Article

PROBABILISTIC ACTUAL CAUSATION

Luke Fenton-Glynn [1,†]

[1] University College London; l.glynn@ucl.ac.uk

† Current address: University College London, Gower Street, London, WC1E 6BT

Academic Editor: name

Version August 1, 2018 submitted to Information

Abstract: Actual (token) causation – the sort of causal relation asserted to hold by claims like *the Chicxulub impact caused the Cretaceous-Paleogene extinction event*, *Mr. Fairchild's exposure to asbestos caused him to suffer mesothelioma*, and *the H7N9 virus outbreak was caused by poultry farmers becoming simultaneously infected by bird and human 'flu strains* – is of significance to scientists, historians, and tort and criminal lawyers. Progress has been made in explicating the actual causal relation in the deterministic case by means of the use of structural equations models and causal graphs. I seek to make similar progress concerning the probabilistic case by using probabilistic causal models and associated causal graphs.

Keywords: probabilistic actual causation; probabilistic causal models; causal graphs

1. Introduction

Actual (token) causation is the sort of causal relation asserted to hold by claims like *the Chicxulub impact caused the Cretaceous-Paleogene extinction event*, *Mr. Fairchild's exposure to asbestos caused him to suffer mesothelioma*,[1] and *the H7N9 virus outbreak was caused by poultry farmers becoming simultaneously infected by bird and human 'flu strains*. Claims about actual causation contrast with generic (type) causal claims, such as *asteroid impacts cause extinction events* and *asbestos exposure causes mesothelioma*. There is little consensus on the relation between actual and generic causation,[2] but my subject matter is actual causation which can be and has been fruitfully theorised about in its own right (classic theories include those of 2, 3, 4, and 5). Actual causation is of significance to scientists, to historians, and to tort and criminal lawyers.

Since the turn of the millennium – and prompted in particular by a hugely influential book written by the computer scientist, Judea Pearl [6] – the formal tools of Structural Equation Models (SEMs) have started to be brought to bear in attempting to provide a clearer understanding of actual causation.[3] SEMs allow for the perspicuous representation of complex patterns of dependence between variables, facilitating the identification of various types of causal structure. This has led to the development of a number of sophisticated accounts of actual causation, enhancing our understanding of that relation.

One limitation of existing accounts, however, is that they mostly focus on the phenomenon of deterministic actual causation (cf. 9, p. 852; 11, p. 498). Roughly speaking, this is the case where the set of all of an event's actual causes suffice for its occurrence. Yet determinism is at best a special case (or even an idealization). According to the orthodox interpretation of quantum mechanics (and several heterodox interpretations), our universe is a non-deterministic, probabilistic one. Moreover, it's not just

---

[1] *Fairchild v Glenhaven Funeral Services Ltd* ([2002] UKHL 22; [2003] 1 AC 32).

[2] My preferred view is that actual causation is the fundamental causal relation, and that generic causal claims simply (generically) quantify over actual causal relations [1].

[3] Such an approach is taken in, for example, [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], and [18].

in fundamental physics that probabilities figure. They also figure heavily in the higher-level sciences, such as statistical mechanics, genetics, ecology, meteorology, and economics. It is thus a significant limitation of a theory of actual causation if it doesn't apply to cases of probabilistic causation, such as an atom's being caused to decay by alpha bombardment, or a person's suffering mesothelioma as a result of asbestos exposure.[4]

In this paper, I will seek to go some way toward addressing this lacuna in a way that builds upon the successes of SEM approaches to deterministic actual causation. Specifically, I will employ the formal tools of probabilistic causal models – probabilistic analogues of deterministic SEMs – and associated causal graphs to provide an analysis of actual causation that is adequate to the probabilistic case.

2. Probability-Raising

Probability-raising is central to the account developed here – as it is on traditional accounts of probabilistic actual causation.[5] To explain how I will understand that notion a bit of stage-setting is required.

Following those who have sought to analyse deterministic actual causation by appeal to SEMs, I take the relata of the actual causal relation to be variable values. Adopting Goldszmidt and Pearl's [26, pp. 669-70] notation, $P(W = w|do(V = v))$ represents the probability of the variable W taking the value $W = w$ that *would* obtain if the variable V were set to the value $V = v$ by an 'intervention' [12, p. 98].[6] This is liable to diverge from the conditional probability $P(W = w|V = v)$: witness the difference between the probability of a storm *conditional* upon the barometer needle pointing toward the word 'storm' and the probability of a storm if I had intervened upon the barometer needle to point it toward 'storm'.

I shall take it that variable X taking the value $X = x$ rather than $X = x'$ raises the probability of variable Y taking the value $Y = y$ in the relevant sense iff:[7]

$$P(Y = y|do(X = x)) > P(Y = y|do(X = x')). \quad (1)$$

Defining probability-raising in this way – rather than via simple conditional probabilities – means avoiding probability-raising relations between independent effects of a common cause, such as the barometer reading and the storm.

The above definition defines a three-place, contrastive, relation of probability-raising: 'X taking the value $X = x$ rather than $X = x'$ raises the probability of $Y = y$'. I take both actual causation and probability-raising to be inherently contrastive.[8] For instance, one might ask: did the driver's driving at 60mph cause the crash? Relatedly: did the driver's driving at 60mph raise the probability of the crash? I suspect the former question is virtually unanswerable. The latter may be answerable, but only if we have a well defined probability distribution over all the possible speeds the driver might have travelled. More tractable, and perfectly cogent, questions are: did the driver's driving at 60mph rather

---

4 The latter depends not only upon the statistical mechanical processes of diffusion, but also (it is believed) upon the molecular bonding processes involved in, for example, the stimulation of macrophages by the HMGB1 protein. These bonding processes are subject to quantum probabilities.

5 See, for example, [19], [20], [21, p. 204], [22], [23, pp. 175-84], [5]. The deficiencies of these accounts have been demonstrated in e.g. [4, pp. 192–202]; [24, pp. 85-96], and [25].

6 The notion of an 'intervention' is itself a causal notion. Still, I follow Woodward [12, pp. 20-2, 104-7] in thinking that an account of causation can be illuminating even if not reductive. If one wishes to pursue a reductive account, then it's worth noting that the Lewisian semantics for counterfactuals Lewis [27] – which does not appeal to causal notions – can be extended to cover the sort of counterfactual appealed to in the present paper (see [28]).

7 Here and throughout, the probabilities should be taken to be those obtaining immediately after all of the interventions bringing about the variable values specified in the scope of the $do(\cdot)$ function have occurred [cf. 23, p. 177].

8 On the contrastivity of actual causation, see [29], [30], [31], [32], and [9, p. 859].

than less than 60mph (alternatively: rather than more than 60mph; or: rather than below the 40mph speed limit) cause the crash? Did the driver's driving at 60mph rather than less than 60mph raise the probability of the crash? The latter questions have two features: (i) they wear their contrastivity on their face; (ii) the contrast is with an alternative relative to which it is (presumably) clear that the driver's driving at 60mph is a probability-raiser and cause of the crash. The former questions were not explicitly contrastive and I take it that the default contrast where no contrast is explicitly stated or made salient by context is that it is simply the complement of the event in question (in this case the driver's driving at 60mph). This is why the questions were difficult to answer: some of the possibilities contained in the complement of the driver's driving at 60mph are possibilities in which she drove at greater than 60mph (and so, presumably, the probability of a crash was still higher) while some of them are possibilities in which she drove at less than 60mph (and so, presumably, the probability of a crash was lower).

Both the formal definition of probability-raising above and the definition of actual causation below are explicitly contrastive. However, since most of the examples that I will discuss can be represented using just sets of binary variables, I shall often suppress mention of the contrast in my informal discussion since it will be obvious that the contrast is just with the only alternative value of the variable in question.[9]

The key problems confronting attempts to analyze probabilistic actual causation – and an informal statement of the solutions that I propose – can be introduced by means of a series of examples, which I describe in the following section.

3. Problem Cases

3.1. Fizzling

The first example is a case of what Schaffer [33, p. 81] calls 'fizzling'.[10] Such cases show that probability-raising – even when understood in the interventionist sense given in the previous section

---

[9] The definitions offered here analyze three-place relations, taking the form 'Variable X taking the value $X = x$ rather than $X = x'$ *raised the probability/was an actual cause* of variable Y's taking the value $Y = y$ iff ...'. But should we in fact take probability-raising and/or actual causation to be a four place relation? Briefly: probability-raising is (as indicated in the main text) naturally taken as a three-place relation, but not as a four-place relation: I don't think 'Variable X taking the value $X = x$ rather than $X = x'$ raised the probability of variable Y taking the value $Y = y$ rather than $Y = y'$' has a clear interpretation. It is more plausible that actual causation should be treated as a four-place relation (cf. [31, p. 328], [12, p. 146]): for example, claims like 'Doctor's administering 20mg rather than 10mg of Drug to Patient caused Patient to recover quickly rather than slowly' seem coherent. A first pass at an analysis of a quaternary relation of actual causation would be: 'Variable X taking the value $X = x$ rather than $X = x'$ was an actual cause of variable Y taking the value $Y = y$ rather than $Y = y'$ iff (i) $P(Y = y|do(X = x)) > P(Y = y|do(X = x'))$ *and* (ii) $P(Y = y'|do(X = x')) > P(Y = y'|do(X = x))$ (e.g. Doctor's administering 20mg rather than 10mg raised the probability of Patient's recovering quickly but Doctor's administering 10mg rather than 20mg would have raised the probability of Patient's recovering slowly). Though I shall not discuss this in detail here, my proposed final analysis of a four-place actual causal relation (in light of the concepts that will be introduced below) would be that variable X taking the value $X = x$ rather than $X = x'$ was an actual cause of variable Y taking the value $Y = y$ rather than $Y = y'$ iff relative to at least one apt model, variable X taking the value $X = x$ rather than $X = x'$ is an unneutralized de facto probability-raiser of variable Y taking the value $Y = y$, while variable X's taking the value $X = x'$ rather than $X = x$ is a de facto probability-raiser of $Y = y'$. Where X and Y are binary variables (as is the case with most of the examples of putative cause and effect variables to be discussed below), it is automatically the case that $X = x'$ (de facto) raises the probability of $Y = y'$ if $X = x$ (de facto) raises the probability of $Y = y$, so we won't need to explicitly take note of this additional complication.

[10] In illustrating each of the causal structures discussed in this section, I will use a vignette that is a variant on one given in [34] (cf. [35, p. 147] and [36, p. 176]), where it is used to illustrate the general phenomenon of probabilistic actual causation. I have deliberately chosen vignettes in which the probabilities involved are quantum, to minimize the scope for dispute over whether they are the sort of objective chances that are apt to underwrite objective actual causal relations. (It is true that the probabilities assigned to agents' making decisions in these examples are not quantum probabilities. But if you're worried that they're therefore not objective, the example could be complicated so that the various 'actions' described depend on the outcomes of quantum measurements rather than ordinary decisions.) In fact, I'm sympathetic to the view that there are objective chances that aren't simply quantum probabilities. (On this latter point, see [37], [38], [39], [40], [41], [42], and [43], *inter alia*).

so as to rule out 'spurious' probability raising between independent effects of a common cause – is not sufficient for causation.

Fizzling

Someone has hooked up a nuclear bomb to a detonator. The detonator is wired to a Geiger counter so that, if the Geiger counter registers above a threshold reading, then the bomb will (with probability ≈ 1) be detonated. I have some Uranium-232 and I place this near to the Geiger counter. The threshold for activating the detonator is such that, given the probabilistic nature of U-232 decay, there's a probability 0.8 that the Geiger counter will register above the threshold.

If I'd refrained from placing the U-232 near to the Geiger counter, then there would have been a 0.00001 chance of the bomb exploding. This is because the bomb is a nuclear device with a Uranium-235 core and there's a certain (very small) probability that enough U-235 atoms in the bomb undergo spontaneous α-decay to start a chain reaction resulting in the bomb's (spontaneously) exploding.

In fact, the threshold reading of the Geiger counter isn't reached, but the bomb nevertheless explodes.

In **Fizzling**, my placing the U-232 near to the Geiger counter was not an actual cause of the bomb's exploding. Rather, the bomb's exploding is a consequence of spontaneous U-235 decay in its core. As a matter of probability, the probabilistic process by which my action threatened to bring about the bomb's explosion 'fizzled out' – as a result of the Geiger threshold not being reached – before it could bring about the explosion.

Suppose that M is a binary variable that takes value $M = 1$ if I place the U-232 near to the Geiger counter and $M = 0$ if I don't. And suppose that E is a binary variable that takes value $E = 1$ if the bomb explodes and $E = 0$ if it does not. Then my action raises the probability of the bomb's exploding in the sense that:

$$P(E = 1|do(M = 1)) = 0.8 > 0.00001 = P(E = 1|do(M = 0)). \quad (2)$$

The example thus shows that probability-raising is not sufficient for actual causation.

Let me briefly anticipate the approach that I shall take below in tackling the problem posed by **Fizzling**. In my view, the probabilistic fact that leads us to consider my action not to be an actual cause of the bomb's exploding in this case is the fact that *the probability of the bomb's exploding given that I placed the U-232 near to the Geiger counter but the Geiger counter did not register above the threshold reading is no higher than the probability of the bomb's exploding would have been if I'd never placed the U-232 near to the Geiger counter in the first place*.

Where T is a binary variable that takes $T = 1$ if the threshold reading is reached and $T = 0$ if it isn't, then this probabilistic fact is expressed formally as follows:

$$P(E = 1|do(M = 1 \& T = 0)) \leq P(E = 1|do(M = 0)). \quad (3)$$

Indeed the terms on the left- and right-hand side of the above inequality are both equal to 0.00001.

In light of the obtaining of the above inequality, we might say that the probability-raising of $E = 1$ by $M = 1$ is 'neutralized' by the fizzling event represented by $T = 0$. This sort of probabilistic 'neutralization' will play an important role in the analysis of probabilistic actual causation to be developed below.

3.2. Probabilistic Preemption

In the previous section we considered an example that shows that probability-raising is not sufficient for actual causation. A second example – a case of probabilistic preemption – illustrates that probability-raising is not necessary either.

Probabilistic Preemption

Someone (neither you nor I) has connected a Geiger counter to a bomb so that the bomb will explode if the Geiger counter registers above a threshold reading. I place a chunk of U-232, which has a half-life of 68.9 years and decays by α-emission, near to the Geiger counter. By chance, enough U-232 atoms decay within a short enough interval for the Geiger counter to reach the threshold reading so that the bomb explodes.

Unbeknownst to me, you've been standing nearby observing. You have a chunk of Th-228, which has a half-life of 1.9 years and decays by α-emission, which contains at least as many atoms as my chunk of U-232. You've decided that you'll place your Th-228 near to the Geiger counter iff I fail to place my U-232 near to the Geiger counter. There's a negligible chance that you won't follow the course of action that you've decided upon. Seeing that I place my U-232 near to the Geiger counter, you don't place your Th-228 near to the Geiger counter.

(The range of α-particles is 3-5 cm. Suppose – here and in the following examples – that, for each of us, to place our chunk 'near' to the Geiger counter is to place it ≤ 5cm away. There is a slight complication if the chance of an alpha particle travelling more than 5cm is not zero, but merely very small, especially if there's a non-zero probability of an alpha particle travelling from wherever your Th-228 is in fact located when you decide not to place it near to the Geiger counter. If there is, then we need to augment the story with the stipulation that no alpha particle traverses the region of space between your Th-228 and the Geiger counter sensor during the relevant interval if we are to conclude that alpha emission from your Th-228 was not an actual cause. If some of the alpha particles that reached the sensor during the relevant interval did in fact *come from your Th-228, and some came from my U-232, then I would take alpha emission from each to be a cause of the bomb's exploding. But for the purposes of the present example, we'll assume that no alpha particles from your Th-228 make it.)*

In this case, my placing my U-232 near the Geiger counter is plausibly regarded as an actual cause of the bomb's exploding. Let M, D, Y, T, and E be binary variables which, respectively, take value 1 if the following things occur (and 0 otherwise): I place my U-232 near to the Geiger counter; you decide to place your Th-228 near to the Geiger counter iff I don't place my U-232 near to the Geiger counter; you place your Th-228 near to the Geiger counter; the threshold Geiger counter reading is reached; the bomb explodes.

Though my act ($M = 1$) was an actual cause of the explosion ($E = 1$), the following inequality holds:

$$P(E = 1|do(M = 1)) < P(E = 1|do(M = 0)). \tag{4}$$

That is, my placing my U-232 near to the Geiger *lowers* the probability of the bomb's exploding. That's because it strongly lowers the probability of your placing your more potent Th-228 near to the Geiger (your Th-228 is more potent because Th-228 has a shorter half life than U-232 and your chunk of Th-228 contains at least as many atoms as my chunk of U-232). Probability-raising is therefore unnecessary for actual causation.

Let me briefly anticipate the approach I shall take in tackling the problem posed by **Probabilistic Preemption** below. In my view, the probabilistic fact that inclines us to consider my action to be an actual cause of the bomb's explosion is the fact that *holding fixed your failure to place your Th-228 near to the Geiger counter, my placing my U-232 near to the Geiger raises the probability of the bomb's exploding*.[11] Formally:

$$P(E = 1|do(M = 1\&Y = 0)) > P(E = 1|do(M = 0\&Y = 0)). \tag{5}$$

[11] Similar views are articulated in [44], [45], and [46].

The idea is that, although my action doesn't straightforwardly raise the probability of the bomb's exploding (because it lowers the probability of your placing your Th-228 near to the Geiger counter), my action does raise the probability of the bomb's exploding *when we hold fixed the background fact that you don't place your Th-228 near to the Geiger counter*. This is a probabilistic analogue of what, in the deterministic case, Stephen Yablo ([47], [48]) has called *de facto dependence*: the latter comprising latent counterfactual dependence that is revealed when certain background facts about the actual world are held fixed. In this context, we might call it *de facto probability-raising*.[12]

The analysis of probabilistic actual causation that will be developed below will incorporate the idea that it is this sort of *de facto* probability-raising, rather than straightforward probability-raising that is a necessary condition for causation.

3.3. Fizzled Preemption

Is de facto probability-raising – that is, probability-raising *holding appropriate background factors fixed* – sufficient for actual causation? Unfortunately, the answer is 'no', as can be seen from the following example, which is a sort of hybrid of the previous two.

Fizzled Preemption

Someone has connected a Geiger counter to a bomb so that the bomb will explode if the Geiger counter registers above a threshold reading. I place a chunk of U-232 near to the Geiger counter. Unbeknownst to me, you've been standing nearby observing. You have a chunk of Th-228, which contains at least as many atoms as my chunk of U-232. You've decided that you'll place your Th-228 near to the Geiger counter iff I fail to place my U-232 near to the Geiger counter. There's a negligible chance that you won't follow the course of action that you've decided on. Seeing that I place my U-232 near to the Geiger counter, you don't place your Th-228 near to the Geiger counter. The threshold reading of the Geiger counter isn't reached, but the bomb nevertheless explodes (there was a very small background chance of enough atoms in its U-235 core spontaneously decaying to start a chain reaction resulting in the bomb's exploding).

This time, my act was not an actual cause of the bomb's exploding because the threshold reading of the Geiger counter was not reached. Nevertheless, exactly the same inequality obtains as in the **Probabilistic Preemption** case, namely:

$$P(E = 1|do(M = 1\&Y = 0)) > P(E = 1|do(M = 0\&Y = 0)) \tag{6}$$

This illustrates that de facto probability-raising is not sufficient for causation.

Just as in **Fizzling**, the probabilistic process via which my action threatens to bring about the explosion is fizzled by the Geiger counter's not reaching the relevant threshold. I suggest that the probabilistic fact that corresponds to the intuition that my action is not an actual cause of the bomb's exploding in this case is that *even holding fixed your (in)action, the probability of the bomb's exploding given my placing my U-232 near to the Geiger counter but the threshold's not being reached is no higher than the probability of the explosion if I had never placed my U-232 near to the Geiger in the first place.* Formally:

$$P(E = 1|do(M = 1\&T = 0\&Y = 0)) \leq P(E = 1|do(M = 0\&Y = 0)) \tag{7}$$

In other words, just as straightforward probability-raising can be neutralized by a fizzling event, so too can de facto probability-raising (which, in light of **Probabilistic Preemption**, is a more plausible *necessary* condition for actual causation than is straightforward probability-raising).

---

[12] One of the challenges to be addressed below is to specify *which* background factors are the appropriate ones to hold fixed.

The upshot is a simple proposal: *probabilistic actual causation is to be identified with unneutralized de facto probability-raising*. It is this proposal that I shall seek to develop in the remainder of the paper. Making the proposal precise will involve a certain amount of stage setting, to be done in the next section.

Before proceeding, it is worth making an observation about the examples introduced so far. This is that, although **Fizzled Preemption** was introduced to make vivid the fact that de facto probability-raising is not sufficient for causation – due to the possibility of its being neutralized –, **Fizzling** and **Probabilistic Preemption** also illustrate this fact. **Fizzling** illustrates it because straightforward probability-raising can be considered to be just a limiting case of de facto probability-raising: it is probability-raising holding the null event fixed. And, in **Fizzling**, although $M = 1$ raised the probability of $E = 1$ holding the null event fixed as per Inequality (2), $M = 1$ was not an actual cause of $E = 1$ because this de facto probability-raising is neutralized as per Inequality (3).

Probabilistic Preemption also illustrates the fact that de facto probability-raising isn't sufficient for actual causation. For, although in **Probabilistic Preemption** my action de facto raises the probability of the explosion *and* is an actual cause of it, your decision ($D = 1$) to place your Th-228 near to the Geiger counter iff I don't also de facto raises the probability of the explosion,[13] even though it is *not* an actual cause of the explosion since you don't place your Th-228 near to the Geiger counter. Indeed, plausibly, your decision straightforwardly raised the probability of the bomb's exploding. That is:

$$P(E = 1|do(D = 1)) > P(E = 1|do(D = 0)) \tag{8}$$

Inequality (8) holds provided there's some chance that $M = 0$, because $D = 1$ raises the probability that the bomb will still explode in the scenario in which $M = 0$.[14] Nevertheless, as we'll see below, this de facto probability-raising (holding the null event fixed) is neutralized by the fact that you in fact don't place your Th-228 near to the Geiger counter.

4. Probabilistic Causal Models

In order to render more precise the idea that actual causation might be identified with unneutralized de facto probability-raising, it is helpful to make use of Probabilistic Causal Models (PCMs). A PCM, $\mathcal{M}$, is a 6-tuple $\langle \mathcal{V}, \mathcal{R}, \mathcal{C}, \Omega, \mathcal{F}, do(\cdot) \rangle$. Here $\mathcal{V}$ is a set of variables. $\mathcal{R}$ is a function from elements of $\mathcal{V}$ to sets of values: for all $V_i \in \mathcal{V}$, $\mathcal{R}(V_i)$ is the *range* of V_i. In Halpern and Pearl's [9, pp. 851-2] terminology, a formula $V_i = v_i$, for $V_i \in \mathcal{V}$ and $v_i \in \mathcal{R}(V)$, is a *primitive event*. $\mathcal{C}$ is the set of all those possible conjunctions of primitive events, $V_1 = v_1 \& \ldots \& V_n = v_n$, such that $V_i \in \mathcal{V}$ and $v_i \in \mathcal{R}(V_i)$ and such that, for no pair of conjuncts $V_i = v_i$, $V_j = v_j$ is $V_i \equiv V_j$, and where no two elements of $\mathcal{C}$ differ *only* in the permutation of their conjuncts. Such a conjunction is denoted $\mathbf{V} = \mathbf{v}$ (primitive events and the null event are limiting cases of such conjunctions). Abusing notation, the fact that $v_i \in \mathcal{R}(V_i)$ for each primitive event $V_i = v_i$ in the conjunction $\mathbf{V} = \mathbf{v}$, is abbreviated $\mathbf{v} \in \mathcal{R}(\mathbf{V})$ and the set of variables that appear in $\mathbf{V} = \mathbf{v}$ is denoted $\mathbf{V}$.

Call a conjunction $\mathbf{V} = \mathbf{v}$ *maximal* iff it contains a conjunct of the form $V_i = v_i$ for each $V_i \in \mathcal{V}$. Ω is the set of all maximal conjunctions of primitive events. $\mathcal{F}$ is a sigma algebra on Ω. Finally, $do(\cdot)$ is a function from elements of $\mathcal{C}$ to probability distributions on $\mathcal{F}$ [cf. 7, pp. 70, 110]: for each element $\mathbf{V} = \mathbf{v}$ of $\mathcal{C}$, $P(\cdot|do(\mathbf{V} = \mathbf{v}))$ is the probability (chance) distribution on $\mathcal{F}$ that *would* obtain if interventions were performed to bring about $\mathbf{V} = \mathbf{v}$.

---

13 At least if we assume that the most probable alternative is for you to decide not to place your Th-228 near to the Geiger counter come what may.

14 That probabilistic preemption cases illustrate both that straightforward probability-raising (i.e. probability-raising under the null contingency) is neither necessary nor sufficient for actual causation was first observed by Menzies ([5], [24]).

The graph $\mathcal{G}$ associated with a PCM, $\mathcal{M} = \langle \mathcal{V}, \mathcal{R}, \mathcal{C}, \Omega, \mathcal{F}, do(\cdot) \rangle$, is an ordered pair $\langle \mathcal{N}, \mathcal{E} \rangle$ such that $\mathcal{N}$ is a set of nodes and $\mathcal{E}$ is a set of directed edges or ordered pairs of nodes, with $\mathcal{N} \equiv \mathcal{V}$ and $\langle V_i, V_j \rangle \in \mathcal{E}$ iff, where $\mathbf{S} = \mathcal{V} \backslash V_i, V_j$, there is some assignment of values $\mathbf{s}' \in \mathcal{R}(\mathbf{S})$, some pair of values $v_i, v_i' \in \mathcal{R}(V_i)$ $(v_i \neq v_i')$ and some value $v_j \in \mathcal{R}(V_j)$ such that $P(V_j = v_j | do(V_i = v_i \& \mathbf{S} = \mathbf{s}')) \neq P(V_j = v_j | do(V_i = v_i' \& \mathbf{S} = \mathbf{s}'))$. The graph $\mathcal{G}$ can be represented diagrammatically by taking the elements of $\mathcal{N}$ as vertices and drawing an arrow from V_i to V_j iff $\langle V_i, V_j \rangle \in \mathcal{E}$.

To construct a PCM, $\mathcal{M}_{\mathbf{Fiz}}$, of **Fizzling**, we might take the variable set $\mathcal{V}_{\mathbf{Fiz}}$ to be $\{M, T, E\}$. The range of each variable in $\mathcal{V}_{\mathbf{Fiz}}$ is the pair $\{0, 1\}$. $\mathcal{C}_{\mathbf{Fiz}}$, $\Omega_{\mathbf{Fiz}}$, and $\mathcal{F}_{\mathbf{Fiz}}$ are generated by $\mathcal{V}_{\mathbf{Fiz}}$ and $\mathcal{R}_{\mathbf{Fiz}}$ in the way described above. For each element of $\mathcal{C}_{\mathbf{Fiz}}$, the function $do(\cdot)$ returns the chance distribution on $\mathcal{F}_{\mathbf{Fiz}}$ that would obtain if interventions were performed to bring about that element of $\mathcal{C}_{\mathbf{Fiz}}$. The graph for $\mathcal{M}_{\mathbf{Fiz}}$ is given as Figure 1.

Figure 1

We can also construct a PCM, $\mathcal{M}_{\mathbf{ProbPre}}$, of **Probabilistic Preemption**. In doing so, we might take the variable set to be $\mathcal{V}_{\mathbf{ProbPre}} = \{D, M, Y, T, E\}$. The range of each variable in $\mathcal{V}_{\mathbf{ProbPre}}$ is the pair $\{0, 1\}$. $\mathcal{C}_{\mathbf{ProbPre}}$, $\Omega_{\mathbf{ProbPre}}$, and $\mathcal{F}_{\mathbf{ProbPre}}$ are generated by $\mathcal{V}_{\mathbf{ProbPre}}$ and $\mathcal{R}_{\mathbf{ProbPre}}$ in the way described above. For each element of $\mathcal{C}_{\mathbf{ProbPre}}$, the function $do(\cdot)$ returns the chance distribution on $\mathcal{F}_{\mathbf{ProbPre}}$ that would obtain if interventions were performed to bring about that element of $\mathcal{C}_{\mathbf{ProbPre}}$. The graph for $\mathcal{M}_{\mathbf{ProbPre}}$ is given as Figure 2.

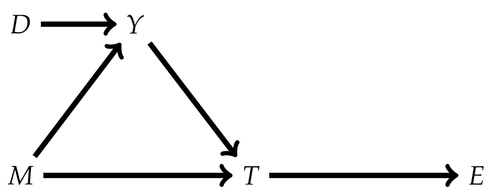

Figure 2

A further piece of terminology will be useful in what follows: a *directed path* in a graph is an ordered sequence of nodes, $\langle V_1, V_2, \ldots, V_n \rangle$, such that there is a directed edge from V_1 to V_2, and a directed edge from V_2 to $\ldots V_n$. $\langle M, Y, T, E \rangle$ is an example of a directed path in the graph of $\mathcal{M}_{\mathbf{ProbPre}}$.

It is worth noting that $\mathcal{M}_{\mathbf{ProbPre}}$ serves just as well as a PCM for **Fizzled Preemption** as it does for **Probabilistic Preemption**. That's because there is no structural difference between the two examples. The difference resides only in the fact that in **Fizzled Preemption**, as a matter of probability, the variable T takes the value $T = 0$ rather than $T = 1$.

In Section 6, I will seek to make precise the idea that actual causation can be identified with unneutralized de facto probability-raising by appealing to PCMs. My specific strategy will be to define a precise notion of unneutralized de facto probability-raising *relative to a PCM*. I will then provide a definition of what it is for $X = x$ (rather than $X = x'$) to count as an actual cause of $Y = y$ *relative to a PCM*: namely this is for $X = x$ (rather than $X = x'$) to bear the relation of unneutralized de facto probability-raising to $Y = y$ *relative to that PCM*. I then define a non-model-relativized notion of actual causation by saying that $X = x$ (rather than $X = x'$) counts as an actual cause of $Y = y$ *simpliciter* provided that $X = x$ (rather than $X = x'$) counts as an actual cause of $Y = y$ relative to at least one *appropriate* PCM. Those who seek to analyze deterministic actual causation in terms of SEMs often adopt a similar strategy of first defining model-relative (deterministic) actual causation, and then defining (deterministic) actual causation *simpliciter* in terms of (deterministic) actual causation relative to at least one appropriate SEM (see, e.g. [10, p. 287], [11, p. 503], [18]).

My strategy requires an account of 'appropriate' models. Such an account is provided in the next section.

5. Appropriate Models

Many of the criteria for an appropriate SEM for evaluating deterministic actual causation that have been advanced in the literature carry over to PCMs, including the following three:

(Partition) For all $V_i \in \mathcal{V}$, the elements of $\mathcal{R}(V_i)$ should form a partition (16, pp. 397-8; 49, p. 182)

(Independence) For no two variables $V_i, V_j \in \mathcal{V}$ ($V_i \neq V_j$) should there be elements $v_i \in \mathcal{R}(V_i)$ and $v_j \in \mathcal{R}(V_j)$ such that the states of affairs represented by $V_i = v_i$ and $V_j = v_j$ are logically or metaphysically related (10, p. 287; 16, p. 397)

(Naturalness) For all $V_i \in \mathcal{V}$, each value in $\mathcal{R}(V_i)$ should represent reasonably natural and intrinsic states of affairs. (cf. 49, 182)

The analysis of actual causation proposed below (like analyses of deterministic actual causation in terms of SEMs) takes all and only values of distinct variables to be potential causal relata. (Partition) insures that we don't thereby miss actual causal relations because they obtain between the values of a single variable. (Independence) insures that we don't mistake stronger-than-causal relations for causal relations. (Naturalness) insures that unnatural or non-intrinsic states of affairs do not get counted as causes and effects (justification of this restriction is given in [23, pp. 190, 263]).[15,16]

A further condition is that should be imposed is that a PCM is appropriate for evaluating whether $X = x$ is an actual cause of $Y = y$ in world θ only if it satisfies (Veridicality):

(Veridicality) For any conjunction $\mathbf{V} = \mathbf{v} \in \mathcal{C}$ taken as an input, the probability distribution $P(\cdot|do(\mathbf{V} = \mathbf{v}))$ yielded as output by the $do(\cdot)$ function of $\mathcal{M}$ should be the objective chance distribution over $\mathcal{F}$ that would$_\theta$ result from interventions setting $\mathbf{V} = \mathbf{v}$. ('Would$_\theta$' indicates that what is required is that this counterfactual be true in θ.)

The idea behind (Veridicality) is simply that, in order to capture genuine causal relations between the variables in its variable set, a PCM must capture how the chance distribution over those variables would truly change under interventions on subsets of those variables. (Veridicality) is an analogue – for PCMs – of the requirement typically imposed that 'appropriate' SEMs encode only true counterfactuals ([10, p. 287], [11, p. 503]).

In the literature in which an analysis of (deterministic) actual causation in terms of SEMs is sought, another condition on model appropriateness is typically added:

(Serious Possibilities) $\mathcal{V}$ should not be such as to generate elements of Ω that represent possibilities "that we consider to be too remote" (10, p. 287; cf. 12, pp. 86-91, 18).

This condition is needed to avoid a variety of counterintuitive results (see, e.g., [10, p. 287]; [12, pp. 86-91]), the canvassing of all of which would take us too far afield. Most straightforwardly, it helps avoid the overgeneration of cases of causation by absence. For instance, JFK's death presumably counterfactually and probabilistically depends upon Paul McCartney's failure to get in the way of the

15 The fact that (Naturalness) refers to 'reasonably' natural and intrinsic states of affairs introduces a certain amount of vagueness into the account, but this is necessary because demanding perfect naturalness would presumably rule out high-level states of affairs as causes and effects, which would be implausible.

16 If *absences* are unnatural states of affairs (cf. 23, pp. 189-93), we might replace (Naturalness) with the requirement that each variable have *at most one value* representing such a state of affairs if we wish to allow (as we will in the following) that absences of reasonably natural events should count as causes and effects and that *events* that are causes and effects may be represented by binary variables that have the absence of these events as one of their values.

oncoming bullets. But most of us would be reluctant to say that the latter is a cause of the former. This is because we view McCartney's getting in the way as a non-serious possibility. (Serious Possibilities) ensures that no appropriate causal model of the events around JFK's death will include a variable representing whether or not he gets in the way. Consequently no such model will treat his failure to do so as a potential cause.

Adopting (Serious Possibilities) results in a certain amount of vagueness and, perhaps, subjectivity. A discussion of whether this is problematic would take us too far afield.[17] But certainly it doesn't put the present account in any worse shape than its deterministic cousins. Moreover, traditional accounts of actual causation – which don't appeal to causal models – also stand in need of appeal to 'serious possibilities' [12, pp. 86-8].[18]

A final requirement – similar to one imposed in the deterministic actual causation/SEM literature – for a model $\mathcal{M}$ to be an appropriate one for evaluating whether $X = x$ (rather than $X = x'$) is an actual cause of $Y = y$ *simpliciter* (i.e. not merely a model-relative actual cause) in world θ is:

(Stability) There is no model $\mathcal{M}^*$ (satisfying Partition, Independence, Naturalness, Veridicality, and Serious Possibilities) with a variable set $\mathcal{V}^*$ such that $\mathcal{V}^* \supseteq \mathcal{V}^{fine}$ – where either $\mathcal{V}^{fine} \equiv \mathcal{V}$ or $\mathcal{V}^{fine}$ differs from $\mathcal{V}$ only in that some of the variables in $\mathcal{V}^{fine}$ are fine-grainings of variables in $\mathcal{V}$ – relative to which $X = x$ (or: the actual value of some fine-graining X^{fine} of X) rather than $X = x'$ (or: rather than each of the values of some fine-graining X^{fine} of X that correspond to $X = x'$) is *not* an actual cause of $Y = y$ (or: the actual value of some fine-graining Y^{fine} of Y). (cf. 16, pp. 394-5; 15; 11, p. 503).

A variable V_1 is a 'fine-graining' of a variable V_2 iff the values of V_2 correspond to equivalence classes of values of V_1 (cf. 50) or, in other words, iff the values of V_1 are realizers of the values of V_2. Identity can be taken as a limiting case of realization so that if some of the values of V_1 are, or represent, states that are identical to (states represented by) values of V_2, the former can still be said to realize the latter. Thus a variable that has a value that represents a car's travelling below 50mph and then values that represent its travelling at 50-60mph, 60-70mph, etc. can be considered a fine-graining of a binary variable that simply represents whether it travels below 50mph or above 50mph.

The idea behind (Stability) is that an appropriate model is a sufficiently rich representation of causal reality that moving to a richer representation by including additional variables representing further factors and/or by representing already-represented factors in more detail would not reveal an apparent actual causal relation to be spurious.[19] This requirement just reflects the obvious fact that we don't want to make causal judgements on the basis of a model that's too impoverished to accurately reflect causal reality. The workings of (Stability) shall be demonstrated on several occasions in the following.

The converse requirement – that a negative verdict about actual causation should not be overturned in a richer model – isn't needed. This is because actual causation (simpliciter) will be defined in terms of actual causation relative to *at least one* appropriate model. A model-relative verdict that $X = x$ (rather than $X = x'$) is not an actual cause of $Y = y$ thus automatically fails to translate into a verdict that $X = x$ (rather than $X = x'$) is not an actual cause (simpliciter) of $Y = y$ if there's a richer (and otherwise appropriate) model relative to which $X = x$ (rather than $X = x'$) is an actual cause of $Y = y$.

---

[17] But see [12, pp. 86–91] on this question.

[18] If one regards appeal to 'serious possibilities' in an account of actual causation as too problematic, one option would be to simply admit that, e.g., McCartney's failure to get in the way was an actual cause of JFK's death. If so, as Blanchard and Schaffer [49, p. 198] point out, (Serious Possibilities) "may be reinterpreted, not as an aptness constraint on models, but as a descriptive psychological claim about which causal models are most readily available to us when we form our causal judgements."

[19] Note that (Stability) thus renders the notion of an appropriate model relative to the causal claim being evaluated.

6. Probabilistic Actual Causation

In this Section, I will define actual causation *simpliciter* in terms of actual causation relative to an appropriate PCM. I will then define model-relative actual causation in terms of a formally precise notion of unneutralized de facto probability-raising relative to a model. I will state the definitions and then show how they work with reference to the three examples that were introduced in Section 3. The definitions are as follows.[20,21]

AC(S)

Where $x, x' \in \mathcal{R}(X)$ and $y \in \mathcal{R}(Y)$, $X = x$ rather than $X = x'$ is an actual cause (simpliciter) *of $Y = y$ in world θ iff $X = x$ and $Y = y$ are the actual values of X and Y in θ and $X = x$ rather than $X = x'$ is an actual cause of $Y = y$ in θ relative to at least one PCM, $\mathcal{M} = \langle \mathcal{V}, \mathcal{R}, \mathcal{C}, \Omega, \mathcal{F}, do(\cdot) \rangle$ (with $X, Y \in \mathcal{V}$), that is appropriate for evaluating whether $X = x$ rather than $X = x'$ is an* actual cause (simpliciter) *of $Y = y$ in θ.*

AC(M-R)

Where $x, x' \in \mathcal{R}(X)$ and $y \in \mathcal{R}(Y)$, $X = x$ rather than $X = x'$ is an actual cause *of $Y = y$ relative to a model $\mathcal{M} = \langle \mathcal{V}, \mathcal{R}, \mathcal{C}, \Omega, \mathcal{F}, do(\cdot) \rangle$ (with $X, Y \in \mathcal{V}$) in world θ iff there is a partition $(\mathbf{Z}, \mathbf{W})$ of $\mathcal{V} \backslash X, Y$ such that, where $\mathbf{W} = \boldsymbol{w}^*$ are the actual values the variables in $\mathbf{W}$ take in θ, for all subsets $\mathbf{Z}'$ of $\mathbf{Z}$ (where $\mathbf{Z}' = \boldsymbol{z}^*$ are the actual values that the variables in $\mathbf{Z}'$ take in θ):*

$$P(Y = y | do(X = x \& \mathbf{W} = \mathbf{w}^* \& \mathbf{Z}' = \mathbf{z}^*)) > P(Y = y | do(X = x' \& \mathbf{W} = \mathbf{w}^*)) \qquad \textbf{(IN)}$$

In **AC(M-R)**, the variables in **W** correspond to the 'background' that is held fixed in looking for *de facto* probability-raising of $Y = y$ by $X = x$ (rather than $X = x'$). Against that background, the variables in **Z** are the potential neutralizers of any such de facto probability-raising.

In an appropriate causal model we can try each and every partition of the variables into background and potential neutralizers in looking for unneutralized de facto probability-raising of $Y = y$ by $X = x$ (rather than $X = x'$). According to **AC(M-R)** the existence of one such partition is all that is needed for model-relative actual causation, and according to **AC(S)** the existence of actual causation relative to at least one appropriate model is all that is needed for actual causation *simpliciter*. What we will see in what follows is that we have the best chance of identifying unneutralized de facto

[20] Those familiar with Halpern and Pearl's ([8], [9]) analyses of deterministic actual causation are invited to see an analogy between their analyses and **AC(M-R)**. **AC(M-R)** was partly inspired by thinking about how a counterpart of Halpern and Pearl's analysis might be developed that is adequate to the probabilistic case. (In fact, there is a disanalogy to Halpern and Pearl's definitions, and an analogy to that given by Halpern in his [51] in that the following analysis requires that the variables in **W** be held fixed at their actual values – a requirement that will be discussed further in Subsection 7.2 and Subsection 7.3 below. Still, the following analysis, unlike that given in Halpern's [51] incorporates a condition that is analogous to Halpern and Pearl's AC2(b). This is indispensible in the probabilistic case because it allows us to deal with fizzling.) I think that the following definitions are adequate to cover deterministic actual causation as the special case where all probabilities are 1 or 0. This is why my definitions take the definiendum to be 'actual cause' rather than 'probabilistic actual cause'.

[21] Note that relativity to a world θ is invoked in three ways in the definition **AC(S)** of actual cause (simpliciter). First, directly, because **AC(S)** requires that $X = x$ and $Y = y$ to be the values X and Y take in θ if they are to be actual cause and effect in θ. Second, indirectly, because **AC(S)** invokes the notion of model-relative actual causation defined by **AC(M-R)**, and **AC(M-R)** takes model-relative actual causation to itself be a world-relative matter. That's because **AC(M-R)** invokes the *actual values* that the variables in **Z** and **W** take in θ. Third, again indirectly, because **AC(S)** invokes the notion of an *appropriate* PCM and the latter is a world-relative notion because whether a model satisfies (Veridicality) (and perhaps (Serious Possibilities)) is world-relative.

probability-raising relative to a model where we allocate to **Z** all the variables that lie on one or more paths from X to Y in the graph associated with that model and where we allocate to **W** all the variables that lie off those paths.

Briefly,[22] the reason for this is that a directed path from X to Y can be through of as representing an actual or potential causal process connecting the two variables. We therefore won't find de facto probability-raising if we hold fixed intermediate variables on all paths between X and Y, since this will serve to 'screen off' the values of Y from those of X. On the other hand, when we find de facto probability-raising along paths from X to Y by holding off-path variables fixed, this is prima facie evidence that the paths represents active causal processes from X to Y. However, these processes might be fizzled, in which case actual values of variables on those paths will serve as neutralizers.

Still, these definitions can best be understood by seeing how they work. Their workings can be illustrated with respect to the three examples introduced in Section 3.

6.1. Fizzling

AC(M-R) counts $M = 1$ as *not* being an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{Fiz}}$ (and the world described in **Fizzling**). To see this, consider the partition of $\mathcal{V}_{\mathbf{Fiz}} \backslash M, E$ in which $\mathbf{W} = \emptyset$ and $\mathbf{Z} = \{T\}$. Relative to this model and this partition, (**IN**) is not satisfied for the subset $\mathbf{Z}' = \{T\}$. That's because Equality (3) holds.

Since there's only one variable in $\mathcal{V}_{\mathbf{Fiz}} \backslash M, E$ – namely, T – there's only one other partition of the variables in $\mathcal{V}_{\mathbf{Fiz}} \backslash M, E$ to consider when looking for unneutralized de facto probability-raising of $E = 1$ by $M = 1$ relative to $\mathcal{M}_{\mathbf{Fiz}}$: namely that in which T is assigned to **W** rather than **Z**. But unneutralized de facto probability-raising isn't revealed by this assignment either. That's because the value of T 'screens off' M from E, so when the value of T is held fixed as part of the assignment $\mathbf{W} = \mathbf{w}^*$, varying the value of M doesn't make a difference to the probability of $E = 1$, and so (**IN**) doesn't hold for the null subset of **Z**.

Given that (**IN**) isn't satisfied for every subset of **Z** relative to any partition of the variables in $\mathcal{V}_{\mathbf{Fiz}} \backslash M, E$, **AC(M-R)** doesn't count $M = 1$ as an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{Fiz}}$. The question is whether **AC(M-R)** counts $M = 1$ as an actual cause of $E = 1$ relative to *any* appropriate model (as is required for **AC(S)** to count $M = 1$ as an actual cause of $E = 1$ *simpliciter*). In showing that it doesn't, it will be helpful to note that a model $\mathcal{M}_{\mathbf{Fiz}}^{T_{fine}}$ that differs from $\mathcal{M}_{\mathbf{Fiz}}$ only by replacing T with a variable T_{fine} that represents the possible Geiger counter readings as finely as possible (i.e. T_{fine} has a value corresponding to each possible precise reading of the Geiger counter) is also not one relative to which **AC(M-R)** counts $M = 1$ as an actual cause of $E = 1$. That's because, just as the actual value of T serves to screen off M from E if held fixed as one of the background variables **W** and neutralizes the (de facto) probability-raising of $E = 1$ by $M = 1$ if included in **Z**, so too does the actual value – $T_{fine} = t^*_{fine}$ – of T_{fine} (which of course represents a reading of the Geiger counter that lies below the threshold reading).

Next, note that $\mathcal{M}_{\mathbf{Fiz}}$ clearly satisfies (Partition) and (Independence). By stipulation, it satisfies (Veridicality) relative to the world described in **Fizzling**. Although (Naturalness) and (Serious Possibilities) are somewhat vague requirements, it seems clear that $\mathcal{M}_{\mathbf{Fiz}}$ isn't a borderline case, but rather clearly satisfies them (certainly it doesn't represent the sort of possibilities or unnatural states of affairs that these requirements are designed to deal with in the deterministic actual causation/SEM literature – cf. 49). $\mathcal{M}_{\mathbf{Fiz}}^{T_{fine}}$ also satisfies each of these conditions. Since $\mathcal{M}_{\mathbf{Fiz}}$ and $\mathcal{M}_{\mathbf{Fiz}}^{T_{fine}}$ satisfy these requirements, no model relative to which **AC(M-R)** is satisfied, but that has a variable set that is a strict subset of $\mathcal{V}_{\mathbf{Fiz}}$ or $\mathcal{M}_{\mathbf{Fiz}}^{T_{fine}}$ will count as 'appropriate', because it will not satisfy (Stability).

But, among models whose variable sets are not strict subsets of $\mathcal{V}_{\mathbf{Fiz}}$ or $\mathcal{M}_{\mathbf{Fiz}}^{T_{fine}}$, is there any appropriate one relative to which model **AC(M-R)** is satisfied? That is, is there any appropriate

[22] More detail on this can be found in (10, 44, and 9, Appendix A.1).

model and partition ($\mathbf{W}$, $\mathbf{Z}$) of that model's variable set (excluding M and E) relative to which (IN) is satisfied for every subset of $\mathbf{Z}$? The answer is 'no'. To see this suppose that a model satisfying (Serious Possibilities), (Partition), (Independence), (Naturalness), (Veridicality) (hereafter abbreviated 'SPINV') includes T (or T_{fine}) in its variable set. Then, since T (T_{fine}) 'screens off' M from E, T (T_{fine}) will have to be included as an element of $\mathbf{Z}$ rather than $\mathbf{W}$ in any partition of the model's variable set if that partition is to reveal de facto probability-raising of $E = 1$ by $M = 1$. Now consider the subset $\mathbf{Z}' = T$ ($\mathbf{Z}' = T_{fine}$) of $\mathbf{Z}$. Because $M = 1$ is only relevant to $E = 1$ because it is relevant to the Geiger counter reading exceeding the threshold, the probability of $E = 1$ when $M = 1$ and T (T_{fine}) takes its actual value $T = 0$ ($T_{fine} = t^*_{fine}$) is no higher than the probability of $E = 1$ when simply $M = 0$, no matter what other variables we might care to include in the model and hold fixed as part of $\mathbf{W}$.

So **AC(M-R)** doesn't count $M = 1$ as an actual cause of $E = 1$ relative to any model with T (or T_{fine}) in its variable set. This means that any otherwise appropriate model (i.e. any model satisfying SPINV) $\mathcal{M}'_{\mathbf{Fiz}}$ relative to which $M = 1$ *is* an actual cause of $E = 1$ can be expanded to a model $\mathcal{M}'^{+T}_{\mathbf{Fiz}}$ (or $\mathcal{M}'^{+T_{fine}}_{\mathbf{Fiz}}$) relative to which $M = 1$ *isn't* an actual cause of $E = 1$ simply by the addition of T (T_{fine}). Provided that $\mathcal{M}'^{+T}_{\mathbf{Fiz}}$ ($\mathcal{M}'^{+T_{fine}}_{\mathbf{Fiz}}$) satisfies SPINV, $\mathcal{M}'_{\mathbf{Fiz}}$ violates (Stability) and is inappropriate. So **AC(S)** will correctly *not* count $M = 1$ as an actual cause *simpliciter* of $E = 1$.

Since the values of T (T_{fine}) form a partition and represent natural states of affairs, (Partition) and (Naturalness) will be satisfied by $\mathcal{M}'^{+T}_{\mathbf{Fiz}}$ (or $\mathcal{M}'^{+T_{fine}}_{\mathbf{Fiz}}$) if they were satisfied by $\mathcal{M}'_{\mathbf{Fiz}}$. Moreover, it doesn't seem that a variable representing possible readings of the Geiger counter in **Fizzling** can represent the sort of non-serious possibilities that (Serious Possibilities) is intended to rule out by those who adopt it.[23] So it doesn't seem that $\mathcal{M}'^{+T}_{\mathbf{Fiz}}$ (or $\mathcal{M}'^{+T_{fine}}_{\mathbf{Fiz}}$) could represent any non-serious possibilities if $\mathcal{M}'_{\mathbf{Fiz}}$ doesn't.

When it comes to (Veridicality), it should be noted that there are multiple ways of expanding a model $\mathcal{M}'_{\mathbf{Fiz}}$ via the addition of T (T_{fine}), each associated with a different $do(\cdot)$ function implying different facts about how the probability distribution over T (T_{fine}) depends upon other variables in the model, and vice versa. In looking for an apt expanded model, we just select the one with the $do(\cdot)$ function that implies the truth (relative to the world we are modelling) about how the objective chance distribution over T (T_{fine}) depends upon other variable in the model, and vice versa.

(Independence) is a little trickier. Might not $\mathcal{M}'_{\mathbf{Fiz}}$ include a variable whose values are logically or metaphysically related to those of T and T_{fine}? There two interesting possibilities.[24] One of these is that $\mathcal{M}'_{\mathbf{Fiz}}$ includes a variable T' that's intermediate in grain between T and T_{fine}.[25] But then $\mathcal{M}'_{\mathbf{Fiz}}$ itself isn't a model relative to which there's unneutralized de facto probability-raising of $E = 1$ by $M = 1$. That's because the actual value of T' (representing a state of affairs that is identical to, or a realizer of, that represented by $T = 0$) screens off M from E and so can't be held fixed as part of $\mathbf{W}$ in

---

23 True, if T_{fine} represents the full range of possible readings of the Geiger counter, there are some possible readings that will have extremely low probability given that I place my U-232 near to the Geiger counter. Still, whenever (Serious Possibilities) has been invoked in the literature it has been to rule out models containing variables corresponding to certain factors that are both non-salient in some causal story and metaphysically distinct from those that are rendered salient (e.g. McCartney getting in the way of the JFK-bound bullets); it is never to rule out models containing variables that represent salient factors in a fine grained way. Nor do I think there's any need to rule such models out. (Indeed, it would strike me as very troubling – especially in light of the discussion of Subsection 7.1 below – if an account of actual causation depended for its success on not representing factors acknowledged to be relevant in too detailed a way.) So I will interpret (Serious Possibilities) as not ruling them out.

24 The third, non-interesting possibility is one where $\mathcal{M}'_{\mathbf{Fiz}}$ includes a variable or variables representing gerrymandered states of affairs that are logically/metaphysically related to those represented by T and T_{fine} – e.g. T_{gerry}, which takes value $T_{gerry} = 1$ if the Geiger counter threshold is reached *or* Trump is US president, and $T_{gerry} = 0$ otherwise. Such a model would straightforwardly violate (Naturalness) and so even if $M = 1$ counted as an actual cause of $E = 1$ relative to it, it could not be inferred that $M = 1$ is an actual cause of $E = 1$ *simpliciter*.

25 Since T is a binary variable, and since T_{fine} represents the Geiger counter reading in the finest possible way, $\mathcal{M}'_{\mathbf{Fiz}}$ couldn't contain a coarse-graining of T or a fine-graining of T_{fine}.

the partition of the variable set of $\mathcal{M}'_{\mathbf{Fiz}}$ if de facto probability-raising is to be revealed. Moreover (if included in $\mathbf{Z}$) the actual value of T' neutralizes any de facto probability-raising of $E = 1$ by $M = 1$.

A second possibility is that $\mathcal{M}'_{\mathbf{Fiz}}$ contains a variable T_{cross}, the values of which are not all realizers of the values of T, but where there is some 'cross-cutting' of them. For instance, it might be that T_{cross} is a ternary variable, with one value representing an interval of Geiger counter readings from 0 to some value n well below the threshold, a second value representing an interval of readings from some value m well above the threshold and upward, and a third value representing the interval of readings (n, m). If the actual reading of the Geiger counter lay in (n, m) then the actual value of T_{cross} wouldn't serve as a neutralizer for any de facto probability-raising of $E = 1$ by $M = 1$ revealed in $\mathcal{M}'_{\mathbf{Fiz}}$ (as it would, for instance, be relative to the partition $\mathbf{W} = \emptyset$, $\mathbf{Z} = \{T_{cross}\}$ of $\mathcal{V}'_{\mathbf{Fiz}} \backslash M, E$ if $\mathcal{V}'_{\mathbf{Fiz}} = \{M, T_{cross}, E\}$), and we couldn't simply add T or T_{fine} to the model to show that the model violates (Stability), since the result would not be a model that satisfied (Independence). However, in such a case, the model *would* violate (Stability) because, since it represents the Geiger counter reading as finely as possible, T_{fine} is a fine-graining of T_{cross} and replacing T_{cross} with T_{fine} would (for reasons explained above) yield a model in which there's no unneutralized de facto probability-raising of $E = 1$ by $M = 1$.

So any model that satisfies SPINV relative to which $M = 1$ counts as an actual cause of $E = 1$ violates (Stability) so that (**AC(S)**) yields the correct result that $M = 1$ is not an actual cause of $E = 1$ *simpliciter*.

6.2. Probabilistic Preemption

6.2.1. $\mathbf{M} = \mathbf{1}$ *is an actual cause (simpliciter) of* $\mathbf{E} = \mathbf{1}$ *in* ***Probabilistic Preemption***

AC(M-R) counts $M = 1$ as an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{ProbPre}}$ (and the world described in **Probabilistic Preemption**). To see this, consider the partition of $\mathcal{V}_{\mathbf{ProbPre}} \backslash M, E$ such that $\mathbf{W} = \{D, Y\}$ and $\mathbf{Z} = \{T\}$. **AC(M-R)** is satisfied because (**IN**) holds for both subsets of $\mathbf{Z}$ ($\emptyset$ and $\{T\}$), as shown by inequalities 9 and 10:

$$P(E = 1|do(M = 1\&D = 1\&Y = 0)) > P(E = 1|do(M = 0\&D = 1\&Y = 0)) \tag{9}$$

$$P(E = 1|do(M = 1\&T = 1\&D = 1\&Y = 0)) > P(E = 1|do(M = 0\&D = 1\&Y = 0)) \tag{10}$$

Inequality 9 indicates that my action raises the probability of the explosion *holding fixed* that (you make your decision but) don't place your Th-228 near to the Geiger counter. The existence of this *de facto* probability-raising reflects the fact that there's a path – $\langle M, T, E\rangle$ – along which $M = 1$ promotes $E = 1$ ($M = 1$ raises the probability of $E = 1$ when we hold fixed the values of all variables off that path).

Inequality 10 indicates that, again holding fixed $D = 1$ and $Y = 0$, the probability of $E = 1$ is higher if I place my U-232 near the Geiger *and the threshold reading is reached* (as is actually the case in **Probabilistic Preemption**) than if I'd not placed my U-232 near to the Geiger counter. This reflects the fact that, not only is there a potential process via which $M = 1$ threatens to bring about $E = 1$, but an important part of that process (namely: the Geiger counter threshold being reached) occurs.

Since **AC(M-R)** implies that $M = 1$ is an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{ProbPre}}$, **AC(S)** yields the (correct) result that $M = 1$ is an actual cause (simpliciter) of $E = 1$ provided that $\mathcal{M}_{\mathbf{ProbPre}}$ is appropriate. $\mathcal{M}_{\mathbf{ProbPre}}$ *is* appropriate. Clearly it satisfies (Partition) and (Independence). It satisfies (Naturalness) because all of the states that its variables represent are reasonably natural. It was stipulated that the $do(\cdot)$ function associated with $\mathcal{M}_{\mathbf{ProbPre}}$ is such that (Veridicality) is satisfied relative to the world described in **Probabilistic Preemption**. Moreover, $\mathcal{M}_{ProbPre}$ doesn't represent the sort of 'non-serious' possibility that (Serious Possibilities) is introduced to rule out (cf. 10; 12, pp. 86-91).

Finally, (Stability) is satisfied because the causal process from my action to the explosion is complete. Holding fixed $D = 1$ and $Y = 0$, the probability of the explosion if $M = 1$ *and* parts of this process occur is higher than the probability of the explosion if simply $M = 0$. Any variable (whose values represent reasonably natural states, form a partition, and are logically and metaphysically independent from the variables in $\mathcal{V}_{\mathbf{ProbPre}}$) that might be added to $\mathcal{V}_{\mathbf{ProbPre}}$ either represents part of this process or it doesn't. If it does, its actual value represents *the occurrence* of part of the process (since in **Probabilistic Preemption** this process is complete). So, if it is added to $\mathcal{V}_{\mathbf{ProbPre}}$, including it in $\mathbf{Z}$ will not prevent (**IN**) from holding for all subsets $\mathbf{Z}'$ of $\mathbf{Z}$. If it doesn't, then adding it to $\mathcal{V}_{\mathbf{ProbPre}}$, including it in $\mathbf{W}$, and holding it fixed at its actual value as part of the assignment $\mathbf{W} = \mathbf{w}^*$ will not make a difference to the fact that (**IN**) holds for all subsets $\mathbf{Z}'$ of $\mathbf{Z}$, since holding fixed $Y = 0$ as part of $\mathbf{W} = \mathbf{w}'$ is already sufficient to ensure this. Either way, there's a partition of the resulting model's variable set relative to which a relation of unneutralized de facto probability-raising of $E = 1$ by $M = 1$ is revealed. **AC(M-R)** requires the existence of only one such partition for $M = 1$ to count as an actual cause of $E = 1$ relative to this richer model. So any such richer model is one relative to which $M = 1$ is an actual cause of $E = 1$, meaning that (Stability) is satisfied by $\mathcal{M}_{\mathbf{ProbPre}}$, and **AC(S)** therefore correctly counts $M = 1$ as an actual cause of $E = 1$ *simpliciter*.

Note that **Probabilistic Preemption** is an *early* preemption case. It could be changed to a case of *late* preemption by supposing that you decide, not to place your Th-228 near to the Geiger counter iff I don't place my U-232 near to the Geiger counter, but rather that you will place your Th-228 near to the Geiger counter iff *the bomb doesn't explode*. Still, in the scenario where I place my U-232 near to the Geiger counter, the threshold is reached and the bomb explodes (and you don't place your Th-228 near to the Geiger counter), my action is an actual cause of the explosion.

We can model this case using the variable set $\mathcal{V}_{\mathbf{LatePre}} = \{M, Y, D', T, E\}$, with M, Y, T, and E given the same interpretations as in $\mathcal{M}_{\mathbf{ProbPre}}$, but with D' given a slightly different interpretation than that given to the variable D in $\mathcal{M}_{\mathbf{ProbPre}}$: specifically D' takes value $D' = 1$ if you decide that you will place your Th-228 near to the Geiger counter iff *the bomb doesn't explode*; and value $D' = 0$ if you don't make that decision. Combined with the veridical $do(\cdot)$ function for this variable set and this scenario, we arrive at a model $\mathcal{M}_{\mathbf{LatePre}}$. But it is easy to verify that the above reasoning that shows that **AC(M-R)** counts $M = 1$ as an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{ProbPre}}$ isn't sensitive to the difference between $\mathcal{M}_{\mathbf{ProbPre}}$ and $\mathcal{M}_{\mathbf{LatePre}}$. And the reasoning that shows that $\mathcal{M}_{\mathbf{ProbPre}}$ is an appropriate model can be carried across to $\mathcal{M}_{\mathbf{LatePre}}$ just by making the relevant substitutions. Consequently **AC(S)** correctly counts $M = 1$ as an actual cause of $E = 1$ in this late preemption scenario. In Subsection 6.2.3 below, we'll see that it also correctly discounts $D' = 1$ as an actual cause of $E = 1$ in this scenario.

6.2.2. $\mathbf{M = 1}$ *isn't an actual cause of* $\mathbf{E = 1}$ *in* ***Fizzled Preemption***

It was observed above that $\mathcal{M}_{\mathbf{ProbPre}}$ also serves as a model for **Fizzled Preemption**. However, in this case we find that the de facto probability-raising of $E = 1$ by $M = 1$ that is revealed by the partition of $\mathcal{V}_{\mathbf{ProbPre}} \backslash M, E$ such that $\mathbf{W} = \{D, Y\}$ and $\mathbf{Z} = \{T\}$ (and indicated by Inequality 9) is neutralized because the actual value of T is in this case $T = 0$ and Inequality 11 holds:

$$P(E = 1 | do(M = 1 \& T = 0 \& D = 1 \& Y = 0)) \leq P(E = 1 | do(M = 0 \& D = 1 \& Y = 0)) \quad (11)$$

Is there an alternative way of partitioning $\mathcal{V}_{\mathbf{ProbPre}} \backslash M, E$ so as to reveal contingent probability-raising that isn't neutralized in **Fizzled Preemption**? (There need only be *one* such partition for **AC(M-R)** to be satisfied.) There isn't. Suppose that we assign T to $\mathbf{W}$ instead of $\mathbf{Z}$. This won't help, since the value of T 'screens off' M from E so (**IN**) would then not hold for the null subset of $\mathbf{Z}$. On the other hand, suppose that we leave T in $\mathbf{Z}$ and reassign D to $\mathbf{Z}$ (we have to leave Y in $\mathbf{W}$ in order to reveal de facto probability-raising of $E = 1$ by $M = 1$). In that case, since $M = 1$ only de facto raises the probability of $E = 1$ because of its relevance to T, the fact that $T = 0$ will continue to neutralize this de facto probability-raising so that (IN) won't hold for the subset $\mathbf{Z}' = \{T\}$ of $\mathbf{Z}$.

So for **Fizzled Preemption**, there's no partition of $\mathcal{V}_{\mathbf{ProbPre}} \backslash M, E$ such that (IN) is satisfied for all subsets of $\mathbf{Z}$ when we consider $M = 1$ as a putative cause of $E = 1$. **AC(M-R)** therefore doesn't count $M = 1$ as an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{ProbPre}}$ and the world described by **Fizzled Preemption**.

But for **AC(S)** to count $M = 1$ as an actual cause of $E = 1$ *simpliciter*, there need only be one appropriate model relative to which **AC(M-R)** counts $M = 1$ as an actual cause of $E = 1$. Is there such a model? There isn't. Suppose a candidate such model includes T. Because M is only relevant to E because of its relevance to T, the value of T 'screens off' the value of M from that of E. This means that, if T is included in $\mathbf{W}$ in the partition $(\mathbf{W}, \mathbf{Z})$ of that model's variable set and held fixed at the value that it actually takes in **Fizzled Preemption** – namely, $T = 0$ – as part of the assignment $\mathbf{W} = \mathbf{w}^*$, then (IN) won't be satisfied for the empty subset of $\mathbf{Z}$. Alternatively, if T is included in $\mathbf{Z}$ then, no matter what other variables are included in the model and assigned to $\mathbf{W}$, (IN) won't be satisfied for the subset $\{T\}$ of $\mathbf{Z}$. This is because $M = 1$ only threatens to bring about $E = 1$ because it threatens to bring about $T = 1$, so no matter what we hold fixed by inclusion on both sides of (IN), the probability of $E = 1$ is no higher if $M = 1$ *and* $T = 0$ than if simply $M = 0$.

So **AC(M-R)** doesn't count $M = 1$ as an actual cause of $E = 1$ relative to any appropriate model with T in its variable set. Note that the same reasoning, *mutatis mutandis*, shows that, equally, **AC(M-R)** doesn't count $M = 1$ as an actual cause of $E = 1$ relative to any appropriate model with T_{fine} in its variable set. This means that any otherwise appropriate model relative to which $M = 1$ *is* an actual cause of $E = 1$ can be expanded to a model in which $M = 1$ *isn't* an actual cause of $E = 1$ simply by the addition of T (T_{fine}). Provided that the expanded model is appropriate, the original model violates (Stability) and so is inappropriate. So **AC(S)** will correctly not count $M = 1$ as an actual cause *simpliciter* of $E = 1$ relative to the world described by **Fizzled Preemption**.

Since the values of T (T_{fine}) form a partition and represent natural states of affairs, (Partition) and (Naturalness) will be satisfied by the expanded model if they were satisfied by the original model. Moreover, it doesn't seem that a variable representing possible readings of the Geiger counter can represent the sort of non-serious possibilities that (Serious Possibilities) is intended to rule out by those who adopt it. With regard to (Veridicality), it should be noted that there are multiple ways of expanding the original model via the addition of T (T_{fine}), each associated with a different $do(\cdot)$ function that entails different facts about how interventions on T (T_{fine}) affect the probability distribution over the remaining variables in the model, and vice versa. In looking for an apt expanded model, we just select the one with the $do(\cdot)$ function that returns the true objective chances in the world described by **Fizzled Preemption**. Where this is so, (Veridicality) will be satisfied but also the actual value of T (T_{fine}) will screen off M from E – if T (T_{fine}) is included in $\mathbf{W}$ – or neutralize any de facto probability-raising of $E = 1$ by $M = 1$ – if T (T_{fine}) is included in $\mathbf{Z}$.

With regard to (Independence), suppose the original model includes a variable T' that's intermediate in grain between T and T_{fine}. Then the original model itself isn't one relative to which there's unneutralized de facto probability-raising of $E = 1$ by $M = 1$. That's because the actual value of T' screens off M from E and so can't be held fixed as part of $\mathbf{W}$ in the partition of the original model's variable set in looking for de facto probability-raising. Moreover (if included in $\mathbf{Z}$) the actual value of T' neutralizes any de facto probability-raising of $E = 1$ by $M = 1$.

A second possibility is that the original model contains a variable T_{cross} the values of which are not all realizers of the values of T but where there is some 'cross-cutting' of them (of the sort described in Subsection 6.1 above). In such a case, we couldn't simply add T or T_{fine} to the model to show that the model violates (Stability) since the result would be a model that doesn't satisfy (Independence). But, in such a case, the model would nevertheless violate (Stability) because, since it represents the Geiger counter reading as finely as possible, T_{fine} is a fine-graining of T_{cross} and replacing T_{cross} with

T_{fine} would yield a model in which there's no unneutralized de facto probability-raising of $E = 1$ by $M = 1$.

So any model that satisfies SPINV relative to which $M = 1$ counts as an actual cause of $E = 1$ violates (Stability) for the world described by **Fizzled Preemption**, so that **AC(S)** yields the correct result that $M = 1$ is not an actual cause of $E = 1$ *simpliciter*.

6.2.3. $\mathbf{D = 1}$ *isn't an actual cause of* $\mathbf{E = 1}$ *in* ***Probabilistic Preemption*** *or* ***Fizzled Preemption***

It is also worth observing that **AC(S)** gives the verdict that $D = 1$ is *not* an actual cause of $E = 1$ relative to either the world described by **Probabilistic Preemption** or the world described by **Fizzled Preemption**. The same considerations show that this is true for both worlds, since the actual value of T – which is the key difference between these worlds – doesn't play any part in the reasoning that shows this to be so. This reasoning is as follows.

For **AC(S)** to count $D = 1$ as an actual cause of $E = 1$ *simpliciter* there must be an appropriate model relative to which **AC(M-R)** counts $D = 1$ as an actual cause of $E = 1$. But there isn't such a model. To see this, first suppose a candidate such model includes Y in its variable set. Because D is only relevant to E because of its relevance to Y, the value of Y 'screens off' the value of D from that of E. This means that, if Y is included in $\mathbf{W}$ in the partition ($\mathbf{W}$, $\mathbf{Z}$) of the model's variable set and held fixed at its actual value ($Y = 0$) as part of the assignment $\mathbf{W} = \mathbf{w}^*$, then (**IN**) won't be satisfied for the empty subset of $\mathbf{Z}$. Alternatively, if Y is included in $\mathbf{Z}$ then, no matter what other variables are included in the model and assigned to $\mathbf{W}$, (**IN**) won't be satisfied for the subset $\{Y\}$ of $\mathbf{Z}$. Specifically, because $D = 1$ only threatens to bring about $E = 1$ because it threatens to bring about $Y = 1$, no matter what we hold fixed by inclusion on both sides of (**IN**) (provided it's not some gerrymandered state of affairs that's metaphysically related to D and/or E – such as *your-not-making-the-decision-or-the-bomb's-exploding* – in which case the model would violate (Naturalness) and (Independence)), the probability of $E = 1$ is no higher if $D = 1$ *and* $Y = 0$ than if simply $D = 0$.

So **AC(M-R)** doesn't count $D = 1$ as an actual cause of $E = 1$ relative to any appropriate model with Y in its variable set. Note that the same reasoning, *mutatis mutandis*, shows that equally **AC(M-R)** doesn't count $M = 1$ as an actual cause of $E = 1$ relative to any appropriate model with Y_{fine} in its variable set, where Y_{fine} includes values representing precisely how far you place your Th-228 from the Geiger (perhaps up to some threshold $>> 5cm$, with a value reserved to represent your not placing your Th-228 within that threshold distance at all). This means that any otherwise appropriate model relative to which $D = 1$ *is* an actual cause of $E = 1$ can be expanded to a model in which $D = 1$ *isn't* an actual cause of $E = 1$ simply by the addition of Y (or Y_{fine}). Provided that the expanded model is appropriate, the original model violates (Stability) and is inappropriate so **AC(S)** will correctly not count $D = 1$ as an actual cause *simpliciter* of $E = 1$.

Since the values of Y (Y_{fine}) form a partition and represent reasonably natural states of affairs, (Partition) and (Naturalness) will be satisfied by the expanded model if they were satisfied by the original model. Neither Y nor Y_{fine} appears to introduce the sort of possibility that (Serious Possibilities) was introduced to rule out by those who endorse it. So it doesn't seem that the expanded model could violate (Serious Possibilities) if the original model doesn't. With regard to (Veridicality), it should be noted that there are multiple ways of expanding the original model via the addition of Y (Y_{fine}), each associated with a different $do(\cdot)$ function that entails different facts about how interventions on Y (Y_{fine}) affect the probability distribution over the remaining variables in the model, and vice versa. In looking for an apt expanded model, we just select the one with the $do(\cdot)$ function that returns the true objective chances in the worlds described by **Probabilistic Preemption** and **Fizzled Preemption** (the chances were identical in each of these stories, so the true $do(\cdot)$ function is the same for both worlds). Relative to such a $do(\cdot)$ function, Y (Y_{fine}) screens off D from E and its actual value neutralizes any de facto probabilitry-raising of $E = 1$ by $D = 1$ that might exist in such a model.

(Independence) is a little trickier. Might not the original model include a variable whose values are logically or metaphysically related to those of Y and Y_{fine}? If that variable represents reasonably natural states (so that the original model satisfies (Naturalness)), then it seems that this variable must simply pertain to what you do (it won't have values that represent, say, *your-placing-your-Th-228-near-to-the-Geiger-counter-or-the-bomb's-exploding*). It might represent what you do in a more fine-grained way: e.g. it might be the variable Y', which takes value $Y' = 0$ if you don't place your Th-228 near to the Geiger counter, $Y' = 1$ if you place it 2.5-5cm from the Geiger counter, and $Y' = 2$ if you place it 0-2.5cm from the Geiger counter. But then it will also be such that its actual value – in this case, $Y' = 0$ – screens off D from E (so that de facto probability-raising of $E = 1$ by $D = 1$ isn't revealed if it is included in **W** in the partition of that model's variable set) and fizzles the process via which your decision threatens to bring about the explosion (so that if it is included in **Z** in the partition of that model's variable set, it serves to neutralize any de facto probability-raising of $E = 1$ by $D = 1$ that might be revealed via that partition). Either way, this means that **AC(M-R)** isn't satisfied in the original model.

Or it might be that the model contains a variable that has values that represent states that cross-cut those represented by Y: e.g. it might be a variable Y_{cross}, which takes value $Y_{cross} = 0$ if you don't place your Th-228 within 7.5 cms of the Geiger counter, $Y_{cross} = 1$ if you place it 2.5-7.5cm from the Geiger counter, and $Y_{cross} = 2$ if you place it 0-2.5cm from the Geiger counter. In that case, its actual value might not act as a neutralizer for $D = 1$ and $E = 1$ (specifically it won't if its actual value is $Y_{cross} = 1$).

In such a case, we couldn't simply add Y or Y_{fine} to the model to show that it violates (Stability) since the result would be a model that doesn't satisfy (Independence). But, in such a case, the model would still in fact violate (Stability) because Y_{fine} is a fine-graining of Y_{cross} and replacing Y_{cross} with Y_{fine} would, for reasons described above, yield a model in which there's no unneutralized de facto dependence.

So any model that satisfies SPINV relative to which $D = 1$ counts as an actual cause of $E = 1$ violates (Stability) for the world described by **Fizzled Preemption**, so that **AC(S)** yields the correct result that $D = 1$ isn't an actual cause of $E = 1$ *simpliciter*.

Now recall our late preemption case, in which you decide that you'll place your Th-228 near to the Geiger counter iff *the bomb doesn't explode* (rather than iff I don't place my U-232 near to the Geiger counter), but in which I place my U-232 near to the Geiger counter, the threshold is reached and the bomb explodes (and you don't place your Th-228 near to the Geiger counter). It was noted that we can model this case using the variable set $\mathcal{V}_{\mathbf{LatePre}} = \{M, Y, D', T, E\}$, with D' being a variable that takes value $D' = 1$ if you decide that you'll place your Th-228 near to the Geiger counter iff *the bomb doesn't explode*; and value $D' = 0$ otherwise. $\mathcal{M}_{\mathbf{LatePre}}$ is a model arrived at by combining this variable set with the veridical $do(\cdot)$ function for this scenario.

Now note that the above reasoning that shows that **AC(M-R)** doesn't count $D = 1$ as an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{ProbPre}}$ isn't sensitive to the difference between $\mathcal{M}_{\mathbf{ProbPre}}$ and $\mathcal{M}_{\mathbf{LatePre}}$, nor is the reasoning that shows that any model satisfying SPINV that does count $D = 1$ as an actual cause of $E = 1$ violates (Stability). Consequently **AC(S)** correctly counts $D' = 1$ as *not being* an actual cause of $E = 1$ in this late preemption scenario. So **AC(S)** handles this late preemption scenario as well as the early preemption scenario.

7. Further Cases

In this section, I will discuss how the present account handles some further cases that have either been raised as posing problems for existing accounts of probabilistic actual causation or that are analogues to cases that have been raised as posing problems for accounts of deterministic actual causation, and so may further test the present account.

7.1. Determinable Effects

A scenario which poses a prima facie challenge for accounts of probabilistic actual causation is given by Hitchcock [25, p. 411]:

Determinable Effect

"Barney smokes, and he also spends a lot of time in the sun. These two proclivities are not connected; for example, Barney is not forced to go outside in order to smoke. Barney's smoking increases the probability that he will get lung cancer. By increasing his probability of getting lung cancer, Barney's smoking increases the overall probability that he will suffer from some form of cancer, and analogously for his exposure to the sun. In fact, Barney develops skin cancer. A fortiori, Barney develops cancer of some form or other."

Let SM be a binary variable that takes the value $SM = 1$ if Barney smokes and $SM = 0$ if he doesn't; let SN be a variable that takes value $SN = 1$ if Barney spends a lot of time in the sun, $SN = 0$ if he doesn't;[26] and let C be a binary variable that takes value $C = 1$ if Barney develops cancer, and $C = 0$ if he doesn't.

Now consider a model $\mathcal{M}_{\mathbf{DetEff}}$ with the variable set $\mathcal{V}_{\mathbf{DetEff}} = \{SM, C\}$ and the the veridical $do(\cdot)$ function for the world described in **Determinable Effect**. Note that $\mathbf{W} = \emptyset$, $\mathbf{Z} = \emptyset$ is the only possible partition of $\mathcal{V}_{\mathbf{DetEff}} \backslash SM, C$ and (**IN**) is satisfied for $\mathcal{M}_{\mathbf{DetEff}}$ and this partition in virtue of the fact that $SM = 1$ straightforwardly raises the probability of $C = 1$. So **AC(M-R)** counts $SM = 1$ as an actual cause of $C = 1$ relative to $\mathcal{M}_{\mathbf{DetEff}}$. Yet $SM = 1$ is not a genuine actual cause of $C = 1$ in **Determinable Effect**. That's because the state represented by $C = 1$ (Barney's developing cancer) is a determinable of which the actual determinate (Barney's developing *skin* cancer) is clearly not an effect of Barney's smoking.

To see that **AC(S)** gives the correct verdict, consider a variable that's a fine-graining of C: namely, C_1, which represents what type of cancer Barney gets, or whether he doesn't get cancer at all. For instance, perhaps $C_1 = 0$ represents his not getting cancer, $C_1 = 1$ represents his developing only lung cancer, $C_1 = 2$ represents his developing only skin cancer, $C_1 = 3$ represents his developing only bowel cancer, ... $C_1 = i$ represents his developing only lung and bowel cancer, and so on. $SM = 1$ doesn't straightforwardly raise the probability of C_1's taking its actual value, $C_1 = 2$. Consequently, $\mathcal{M}_{\mathbf{DetEff}}$ violates (Stability). Nor is there any reasonably natural state of affairs we could hold fixed to reveal de facto probability-raising of $C_1 = 2$ by $SM = 1$. This means that any appropriate model containing C_1 isn't one relative to which there's de facto probability-raising of $C_1 = 2$ by $SM = 1$, and so isn't one relative to which **AC(M-R)** counts $SM = 1$ is an actual cause of the actual value taken by C_1. In consequence, any model that contains C rather than C_1 relative to which there is de facto probability-raising of $C = 1$ by $SM = 1$ violates (Stability). So **AC(S)** correctly doesn't count $SM = 1$ as an actual cause of $C = 1$.

7.2. Overdetermination of Probabilities

AC(M-R) requires that the variables **W** be held fixed at their actual values in seeking de facto probability-raising and therefore model-relative actual causation. In this respect, it's more analogous to the analysis of deterministic actual causation given in [51] than that given in [8] and [9].

Is there a case for relaxing this requirement? The principal motive for considering a relaxation comes from cases where the probability of an effect is *overdetermined*. Cases where an *effect* is overdetermined – oft discussed in the literature on deterministic causation (the *locus classicus* being [23, pp. 207-12]) – are the special case in which the probability of the effect is, as it were, overdetermined to be 1.[27] Deterministic preemption cases are a species of overdetermination, but here what I have in

---

[26] We could have let SM be a more *fine grained* variable representing *how much* Barney smokes, and SN be a fine-grained variable representing *how much* time he spends in the sun. This would not substantively affect the analysis, except that we'd need to explicitly state the contrasts in evaluating actual causal claims relative to the example (e.g. 'Barney's spending an average of 1000 hours per year in the sun *rather than fewer* was an actual cause of his developing cancer').

[27] Hitchcock [11, p. 512] takes the view that 'trumping' cases – as described by Schaffer [52] – are simply overdetermination cases. I agree and I think that attempts to give probabilistic variants of trumping scenarios will simply issue in cases like

mind are cases where it would be intuitively wrong to say that some of the overdeterminers are causes whereas others are not (in preemption cases the premptor is a cause, while the preempted backup isn't a cause). This sort of case is often referred to as 'symmetric overdetermination' in the deterministic causation literature.[28]

The following is a case of overdetermination of probabilities where those probabilities aren't 1s or 0s. I don't claim that the example is physically realistic, but it serves to illustrate a type of probabilistic causal structure that seems possible in principle.

Overdetermination of Probabilities

Someone (neither you nor I) has connected a Geiger counter to a bomb so that the bomb will explode if the Geiger counter registers above a threshold reading. I place a chunk of U-232 near to the Geiger counter at the same time as you place an identical chunk of U-232 (one comprising exactly the same number of atoms) at an identical distance. The chance of the threshold being reached if either of us alone had placed our U-232 chunks near to the Geiger counter would have been 0.8. Now suppose that, because of their electric charge, alpha particles approaching a location (specifically: the Geiger counter sensor) from two different sources have a propensity to deflect one another from that location. And suppose that the interference between the alpha particles emitted by the two sources is such as to exactly offset the doubled quantity of radioactive material that is present (compared with a scenario in which one of us alone had placed our U-232 near to the Geiger counter), so that the chance of the threshold reading being reached given that we both place our U-232 chunks is identical to the chance of its being reached if either of us had alone placed our U-232 near to the Geiger counter: namely, 0.8. Let us additionally suppose (for simplicity) that each of us was antecedently certain to place our U-232 near to the Geiger counter. Finally, let us suppose that in fact enough non-deflected alpha particles – including some from each source – arrive at the Geiger sensor that the threshold reading is reached and the bomb explodes.

It seems definitely wrong to say that, in **Overdetermination of Probabilities**, one of our actions was an actual cause of the bomb's exploding, while the other's wasn't. However, neither action raises the probability of the bomb's exploding. Specifically, where E, M, and Y have the usual interpretations, the following equality holds:[29]

$$P(E=1|do(M=1)) = P(E=1|do(M=0)) = P(E=1|do(Y=1)) = P(E=1|do(Y=0)) = 0.8 \tag{12}$$

Note also that, in this case, de facto probability-raising of $E=1$ by $M=1$ can't be revealed by holding fixed Y at it's actual value, $Y=1$. Nor can de facto probability-raising of $E=1$ by $Y=1$ be revealed by holding fixed M at its actual value $M=1$. Specifically, the following two equalities hold:

$$P(E=1|do(M=1\&Y=1)) = P(E=1|do(M=0\&Y=1)) = 0.8 \tag{13}$$

---

those described in this section, or in cases in which multiple causes raise the probability of their effects (and therefore de facto do so) in a way that isn't neutralized. So I don't think a discussion of these cases would add anything interesting.

[28] I prefer to try to avoid this terminology, however, because there are cases of the phenomenon that aren't particularly symmetric. For instance, Assassin B decides she'll shoot at Victim iff Assassin A does so. Assassin A is located much further away from Victim than is Assassin B. Seeing the puff of smoke from A's gun, B shoots at Victim. Because of the shorter distance between them, B's bullet pierces Victim's heart simultaneous with A's bullet. A and B's shots overdetermine Victim's death and it seems wrong to say that one is a cause and the other isn't. But there are clear asymmetries in that B's shot occurs later than, and depends upon, A's shot.

[29] Although the example has been constructed so that our actions individually make no difference to the probability (the probability is 'overdetermined'), it could easily be transformed into an example in which each of our acts lowers the probability of the explosion by supposing the interference between the sources is stronger still. Making this alternative assumption makes no difference to the way in which **AC(M-R)** and **AC(S)** diagnose the case.

$$P(E=1|do(Y=1\&M=1)) = P(E=1|do(Y=0\&M=1)) = 0.8 \quad (14)$$

In general, it seems the only way that we could reveal a latent probability-raising of $E = 1$ by $M = 1$ is by holding fixed Y at its non-actual value $Y = 0$, and likewise the only way that we could reveal latent probability-raising of $E = 1$ by $Y = 1$ is by holding fixed M at its non-actual value $M = 0$. Thus (assuming that the probability of the bomb's spontaneously exploding is 0.00001):

$$P(E=1|do(M=1\&Y=0)) = 0.8 > 0.00001 = P(E=1|do(M=0\&Y=0)) \quad (15)$$

$$P(E=1|do(Y=1\&M=0)) = 0.8 > 0.00001 = P(E=1|do(Y=0\&M=0)) \quad (16)$$

So we might consider whether we wish to modify **AC(M-R)** so that the setting of **W** isn't required to be the actual values of **W**.

But whether we wish to do so or not very much depends upon our attitude toward overdetermination (of probabilities). I'm inclined to agree with Lewis [23, pp. 207, 212] that it's intuitively doubtful that overdeterminers are causes. *Between them* the overdeterminers cause the outcome [23, pp. 207, 212], but I think it's doubtful that individually they are causes. In this vein Halpern [51, p. 3026] says: "Intuitively, ... they [the overdeterminers] are both parts of a cause, but not causes." The reason I think it's doubtful that individual overdeterminers are causes is precisely that, in the circumstances, they don't make a difference to whether or not (or with what probability) the effect occurs [cf. 51].[30]

For those sympathetic to so called 'difference-making' approaches to actual causation (and note that SEM and PCM accounts of actual causation are among such approaches) there is the following reason for doubting that overdeterminers are causes. We all know that there are things that are plainly actual causes that don't straightforwardly make a difference to their effects (either by making a difference to whether-or-not they occur or by making a difference to the probability that they occur). Preemption cases show this.[31] The intuition that motivates the de facto dependence/de facto probability-raising approach is that, nevertheless, a latent counterfactual dependence or probability-raising can be recovered by holding fixed certain features of the actual situation. But what would motivate analyzing actual causation in terms of a 'latent' probability-raising that is only present against a non-actual background. Why should the existence of difference-making holding fixed factors that are only present in a non-actual possible world indicate the existence of actual causation here in this world?

Can we afford not to take a stand on the present issue? That is, could we simply say that those who think each of a set of overdeterminers (of probabilities) is itself a cause should feel welcome to adopt an analysis of actual causation that weakens **AC(M-R)** by not requiring that the setting of **W** be to the actual values of **W**, while those of us who think that elements of a set of overdeterminers

---

30 Sometimes overdeterminers might make a difference to the timing or manner in which an effect occurs. That is, they make a difference to whether or not (or the probability of whether) a certain 'alteration' [53] of the effect occurs. If so, they are not overdeterminers of the alteration. If we decide to identify the effect with the alteration (i.e. to take the effect to be modally fragile), then we just have a regular case where an effect has two causes (that are not overdeterminers) and our unmodified analysis will treat it as such. If we decide to identify the effect with the original, non-modally fragile effect, then there's genuine overdetermination (of probabilities). Unlike Lewis [53], I don't think that influencing which alteration of an event occurs is necessary or sufficient for causation of that event. For a demonstration of why it is not, see [54].

31 **Probabilistic Preemption** in fact shows that actual causes may *lower* the probability of their effects. But it can be modified into a case where the preempting cause makes no difference to the probability of its effect. For this to be the case, what we would need is for the negative relevance of my act to the explosion that comes by way of its negative relevance to your potential act to exactly balance (rather than 'overbalance') my act's tendency to promote the explosion because of the possibility that my U-232 will emit enough alpha particles that the Geiger counter reaches the threshold reading. This could be achieved by supposing that my act less strongly lowers the probability of your placing your radioactive material near to the Geiger counter than in the original example and/or supposing that your radioactive material is less potent than it is in the original example.

(of probabilities) are not actual causes should stick with the original version of **AC(M-R)**? (Perhaps different people just have slightly different concepts of actual causation.) In fact, weakening the requirement that the setting of **W** be to the actual values of **W** comes at a cost, as we shall see in Subsection 7.3 below. This cost is that there seem to be cases of clear non-causation where it is possible to reveal a 'latent' (and unneutralized) probability-raising by holding fixed certain variables at non-actual values.

Before showing this, it is worth noting that one modification to **AC(M-R)** and **AC(S)** *is* needed if these definitions are to capture the whole of the view of overdetermination cases that I have advocated. Specifically, while these definitions already capture the idea that individual overdeterminers aren't actual causes, they don't capture the view that, taken together, they are an actual cause. To capture this, we just need to modify the definitions by adopting a familiar feature of SEM approaches to actual causation (one possessed by the definitions given in 9, 17, and 51): namely, to allow that an actual cause might be a vector of values, rather than a single value. That is to say, in **AC(M-R)** and **AC(S)**, we would replace '$X = x$' with '$\mathbf{X} = \mathbf{x}$' and '$X = x'$' with '$\mathbf{X} = \mathbf{x}'$' throughout. In **AC(M-R)**, we'd replace '... there is a partition $(\mathbf{Z}, \mathbf{W})$ of $\mathcal{V} \backslash X, Y$ ...' with '... there is a partition $(\mathbf{Z}, \mathbf{W})$ of $\mathcal{V} \backslash \mathbf{X}, \mathbf{Y}$ ...'.

Once we made this modification we would also need to introduce a 'minimality' condition into our analysis (as is also introduced in [9], [17], and [51]) in order to ensure that events that are intuitively not parts of causes don't get counted as such. This would simply involve adding to **AC(M-R)** the requirement that there is no strict subset $\mathbf{X}^{\subset}$ of $\mathbf{X}$ such that, where $\mathbf{X}^{\subset} = \mathbf{x}^{\subset}$ denotes the variables in $\mathbf{X}^{\subset}$ taking the values that they receive in the assignment $\mathbf{X} = \mathbf{x}$, and $\mathbf{X}^{\subset} = \mathbf{x}^{\subset\prime}$ denotes the variables in $\mathbf{X}^{\subset}$ taking the values that they receive in the assignment $\mathbf{X} = \mathbf{x}'$, **AC(M-R)** (or rather the rest of it, once this condition is added) is satisfied when $\mathbf{X}$, $\mathbf{X} = \mathbf{x}$ and $\mathbf{X} = \mathbf{x}'$ are respectively replaced by $\mathbf{X}^{\subset}$, $\mathbf{X}^{\subset} = \mathbf{x}^{\subset}$, and $\mathbf{X}^{\subset} = \mathbf{x}^{\subset\prime}$ throughout.

This yields the desired result in the present case. To see this, consider the simple model $\mathcal{M}_{\mathbf{Overdet}}$ generated from the variable set $\mathcal{V}_{\mathbf{Overdet}} = \{M, Y, E\}$ and the veridical $do(\cdot)$ function for the world described in **Overdetermination of Probabilities**. Let $\mathbf{X} = \{M, Y\}$, so that $\mathbf{W}, \mathbf{Z} = \emptyset$ in the partition of $\mathcal{V}_{\mathbf{Overdet}} \backslash \{M, Y\}, E$ and note that the variables in $\mathbf{X}$ taking the values $M = 1 \& Y = 1$ rather than $M = 0 \& Y = 0$ raises the probability of $E = 1$ when we hold fixed the variables in $\mathbf{W}$ (of which there are none) at their actual values. That is:

$$P(E = 1 | do(M = 1 \& Y = 1)) > P(E = 1 | do(M = 0 \& Y = 0)) \quad (17)$$

Thus the conjunctive event $M = 1 \& Y = 1$ de facto raises the probability of $E = 1$ relative to an alternative in which $M = 0 \& Y = 0$. Since $\mathbf{Z} = \emptyset$, this contingent probability-raising is automatically not neutralized. Note, moreover, that the minimality condition is satisfied, since neither $M = 1$ nor $Y = 1$ straightforwardly or de facto raises the probability of $E = 1$, as equalities 12, 13, and 14 indicate. (So neither $M = 1$ nor $Y = 1$ is alone an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{Overdet}}$.) This means that **AC(M-R)** counts $M = 1 \& Y = 1$ rather than $M = 0 \& Y = 0$ as an actual cause of $E = 1$ relative to $\mathcal{M}_{\mathbf{Overdet}}$.

The model $\mathcal{V}_{\mathbf{Overdet}} = \{M, Y, E\}$ plainly satisfies (Partition), (Independence), and (Naturalness). It doesn't represent any non-serious possibilities relative to the world described in **Overdetermination of Probabilities**, so it also satisfies (Serious Possibilities). Moreover, by assumption, the processes connecting your action and my action to the bomb's exploding is complete (enough alpha particles from our U-232 chunks reach the detector). Consequently, a richer (in the sense of containing more variables) but otherwise appropriate model will not be one in which this de facto probability-raising is neutralized. Nor would any fine-graining of the variables involved eliminate this unneutralized de facto probability-raising. So (Stability) is also satisfied.

AC(S) therefore yields the verdict that $M = 1 \& Y = 1$ rather than $M = 0 \& Y = 0$ is an actual cause of $E = 1$, even though neither $M = 1$ nor $Y = 1$ is individually an actual cause of $E = 1$. This seems to me to be exactly the right result.

7.3. Overlappings

The following example, due to Schaffer [55, p. 41] (a similar example is given by Hitchcock [25, pp. 415-16]), further illustrates why we ought to be sceptical that 'de facto' probability-raising that is revealed only by holding variables fixed at non-actual values should be taken to be indicative of actual causation:

Overlapping

"An atom of U-238 and an atom of Ra-226 are placed in a box at t0 (assume for simplicity that the box is otherwise empty). At t1 the box contains an atom of Th-234, an alpha particle, and (still) an atom of Ra-226. The relevant physical laws are: (1) an atom of U-238 has a certain chance per unit interval of producing Th-234 and an alpha particle, (2) an atom of Ra-226 has a certain chance per unit interval of producing Rn-222 and an alpha particle, and (3) these chances are independent. ... And note that the Ra-226 aims to produce the alpha particle directly, since particle emissions from radioactive sources (as standardly understood) occur without hidden intermediaries".

It's clear, in this example, that the t_0 presence of the Ra-226 atom isn't an actual cause of the t_1 presence of the alpha particle, though the former straightforwardly raises the probability of the latter. The case has a different structure to **Fizzling**, since there's no intermediate process by which Ra-226 produces alpha particles, and so no process to be incomplete or fizzled.

Still, this case can be handled by the present account, since the (continued) presence of the Ra-226 atom at t_1 is itself a neutralizer for the de facto probability-raising of the t_1 presence of the alpha particle by the t_0 presence of the Ra-226 atom.

To see this, consider a model $\mathcal{M}_{\textbf{Overlap}}$ comprising a binary variable R_{t_0} that takes the value $R_{t_0} = 1$ if the Ra-226 atom is present at t_0 and $R_{t_0} = 0$ otherwise, a binary variable R_{t_1} that takes the value $R_{t_1} = 1$ if the Ra-226 atom is present at t_1 and $R_{t_1} = 0$ otherwise, a binary variable U_{t_0} that takes the value $U_{t_0} = 1$ if the U-238 atom is present at t_0 and $U_{t_0} = 0$ otherwise, a binary variable U_{t_1} that takes the value $U_{t_1} = 1$ if the U-238 atom is present at t_1 and $U_{t_1} = 0$ otherwise, a binary variable A_{t_0} that takes the value $A_{t_0} = 1$ if there's an alpha particle present at t_0 and $A_{t_0} = 0$ otherwise, and a binary variable A_{t_1} that takes the value $A_{t_1} = 1$ if there's an alpha particle present at t_1 and $A_{t_1} = 0$ otherwise. And suppose that the $do(\cdot)$ function associated with $\mathcal{M}_{\textbf{Overlap}}$ is the veridical one for the world described by **Overlapping**.

Consider the partition of $\mathcal{V}_{\textbf{Overlap}} \backslash R_{t_0}, A_{t_1}$ such that $\mathbf{W} = \{A_{t_0}, U_{t_0}, U_{t_1}\}$ and $\mathbf{Z} = \{R_{t_1}\}$. Now, if there were no chance either of the U-238 or of an alpha particle produced by the U-238 decay escaping the box, then (since U-238 can only decay via the emission of an alpha particle) there's not even de facto probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$ relative to this partition, since in that case:

$$P(A_{t_1} = 1 | do(R_{t_0} = 1 \& A_{t_0} = 0, U_{t_0} = 1 \& U_{t_1} = 0)) = 1 = P(A_{t_1} = 1 | do(R_{t_0} = 0 \& A_{t_0} = 0 \& U_{t_0} = 1 \& U_{t_1} = 0)) \quad (18)$$

That's because, if the U-238 is in the box at t_0 but not at t_1 and if there's zero chance of either the U-238 or an alpha particle escaping the box then the U-238 must have decayed and the alpha-particle must be in the box.

But let's suppose that there's some chance of either the U-238 or an alpha particle that it produces escaping the box between t_0 and t_1. Then the t_1 presence of the Ra-226 raises the probability of there being an alpha particle in the box at t_1 when the variables in $\mathbf{W}$ are held fixed at their actual values, $\{A_{t_0} = 0, U_{t_0} = 1, U_{t_1} = 0\}$. That is:

$$P(A_{t_1} = 1|do(R_{t_0} = 1\&A_{t_0} = 0\&U_{t_0} = 1\&U_{t_1} = 0)) > P(A_{t_1} = 1|do(R_{t_0} = 0\&A_{t_0} = 0\&U_{t_0} = 1\&U_{t_1} = 0)) \tag{19}$$

That is: the presence of the Ra-226 at t_0 raises the probability of the presence of an alpha-particle in the box at t_1, even given the t_0 presence and t_1 absence of the U-238 (and the t_0 absence of an alpha particle), because there's some chance that the U-238 simply escapes the box or that an alpha particle that it produces does so and the t_0 presence of the Ra-226 raises the probability that there will nevertheless be an alpha particle present in the box at t_1 in such circumstances.

However, the de facto probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$ is neutralized. That's because:

$$P(A_{t_1} = 1|do(R_{t_0} = 1\&R_{t_1} = 1\&A_{t_0} = 0\&U_{t_0} = 1, U_{t_1} = 0)) = P(A_{t_1} = 1|do(R_{t_0} = 0\&A_{t_0} = 0\&U_{t_0} = 1, U_{t_1} = 0)) \tag{20}$$

That is, (in circumstances in which no alpha particle is present at t_0 and the U-238 atom is present at t_0 but absent at t_1) the probability that an alpha particle is present at t_1 if the Ra-226 atom is present both at t_0 and t_1 is no higher than if the Ra-226 hadn't even been present at t_0.

Of course, for **AC(M-R)** to count $R_{t_0} = 1$ as an actual cause of $A_{t_1} = 1$ relative to $\mathcal{M}_{\mathbf{Overlap}}$ it need only be the case that there's *one* partition of $\mathcal{V}_{\mathbf{Overlap}} \backslash R_{t_0}, A_{t_1}$ relative to which $R_{t_0} = 1$ bears the relation of non-neutralized de facto probability-raising to $A_{t_1} = 1$. But brief reflection makes it clear that this is not the case.

To see this, note that assigning A_{t_0}, U_{t_0} or U_{t_1} to **Z** rather than **W** while leaving R_{t_1} in **Z** won't help. No matter whether we hold fixed one, two, all, or none of A_{t_0}, U_{t_0} and U_{t_1} at their actual values, the probability of an alpha particle's being present at t_1 given that the Ra-226 was present at t_0 and at t_1 is no higher than if the Ra-226 had never been present in the first place. Thus the de facto probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$ is neutralized for the subset $\{R_{t_1}\}$ of **Z**. Nor will assigning R_{t_1} to **W** rather than **Z** help (no matter what combination of A_{t_0}, U_{t_0}, U_{t_1} or none is left in **W**). That's because the actual value $R_{t_1} = 1$ of R_{t_1} serves to 'screen off' the value of R_{t_0} from that of A_{t_1} so, if we include $R_{t_1} = 1$ as part of the fixed background $\mathbf{W} = \mathbf{w}^*$, then no relation of de facto probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$ will be revealed. Consequently **AC(M-R)** doesn't count $R_{t_0} = 1$ as an actual cause of $A_{t_1} = 1$ relative to $\mathcal{M}_{\mathbf{Overlap}}$.

Note that this case illustrates why it's important that we not allow non-actual settings of the variables in **W**. That's because if we let $\mathbf{W} = \{A_{t_0}, U_{t_0}, U_{t_1}, R_{t_1}\}$ and $\mathbf{Z} = \emptyset$, then 'de facto' probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$ would be revealed by holding fixed the variables in **W** at the non-actual assignment $\{A_{t_0} = 0, U_{t_0} = 0, U_{t_1} = 0, R_{t_1} = 0\}$ or the non-actual assignment $\{A_{t_0} = 0, U_{t_0} = 1, U_{t_1} = 1, R_{t_1} = 0$. That is, in (non-actual) circumstances in which there's no alpha particle present at t_0 and in which the U-238 is present at both t_0 and t_1 or absent at both t_0 and t_1 *and* in which the Ra-226 is absent at t_1, there's more chance of an alpha particle's being present at t_1 if the Ra-226 is present at t_0. That's because the t_0 presence of the Ra-226 makes it more likely that the t_1 absence of the Ra-226 is due to decay rather than the Ra-226 just having not being in the box in the first place. Note that, since $\mathbf{Z} = \emptyset$ in the partition of $\mathcal{V}_{\mathbf{Overlap}}$ presently under consideration, this 'de facto' probability-raising is automatically not neutralized. So if **AC(M-R)** allowed non-actual settings of **W** to be considered, it would yield the result that $R_{t_0} = 1$ is an actual cause of $A_{t_1} = 1$ relative to $\mathcal{M}_{\mathbf{Overlap}}$. Moreover, $\mathcal{M}_{\mathbf{Overlap}}$ appears to be an appropriate model. (Naturalness), (Independence), (Partition), and (Serious Possibilities) all appear to be satisfied. By stipulation, (Veridicality) is satisfied relative to the world described in **Overlapping**. Finally, it's not at all obvious how $\mathcal{M}_{\mathbf{Overlap}}$ could be expanded or refined into an appropriate model relative to which $R_{t_0} = 1$ doesn't 'de facto' raise the probability of $A_{t_1} = 1$, so it seems plausible that (Stability) is satisfied. So if **AC(M-R)** allowed non-actual settings of **W** to be considered, it appears that **AC(S)** would yield the erroneous result that $R_{t_0} = 1$ is an actual cause (simpliciter) of $A_{t_1} = 1$.

The disallowance of non-actual settings of **W** is no *ad hoc* manoeuvre. Clearly the central reason that we identify the t_0 presence of the Ra-226 atom as *not* being an actual cause of the t_1 presence of the alpha particle is precisely that the Ra-226 atom is still present at t_1. Considering what would have been the case had the Ra-226 atom not been present at t_1 seems clearly irrelevant and misleading regarding the causal status of the t_0 presence of the Ra-226 atom.

Because of its disallowance of non-actual settings of **W**, **AC(M-R)** doesn't count $R_{t_0} = 1$ as an actual cause of $A_{t_1} = 1$ relative to $\mathcal{M}_{\textbf{Overlap}}$. Indeed, **AC(M-R)** doesn't count $R_{t_0} = 1$ as an actual cause of $A_{t_1} = 1$ relative to any otherwise appropriate model that includes R_{t_1} in its variable set. That's because, in the partition $\{\mathbf{W}, \mathbf{Z}\}$ of the model's variable set, either R_{t_1} will be included in **W** in which case no de facto probability raising of $A_{t_1} = 1$ by $R_{t_0} = 1$ is revealed because the actual value of R_{t_1} screens off R_{t_0} from A_{t_1}, or R_{t_1} is included in **Z** where its actual value serves to neutralize any de facto probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$. This means that any otherwise appropriate model relative to which there's unneutralized de facto probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$ can be expanded into one relative to which there isn't by means of the addition of R_{t_1} to its variable set, meaning that the original model violates (Stability). (What if the original model contains a variable that's logically or metaphysically related to R_{t_1}? Well, if it satisfies (Naturalness), it seems that this could only be by containing a variable that's a fine-graining of R_{t_1}, since there doesn't seem to be any way of cross-cutting the states represented by the values of R_{t_1} while still representing reasonably natural states of affairs. On the other hand, a fine-graining of R_{t_1} might, for instance, represent the position and velocity of the Ra-226 atom if present. But the actual value of such a variable – which would represent the Ra-226 being present, perhaps at such-and-such a location in the box and with so-and-so a velocity – would itself screen off R_{t_0} from A_{t_1} and neutralize any de facto probability-raising of $A_{t_1} = 1$ by $R_{t_0} = 1$. So the latter wouldn't count as a cause of the former relative to such a model anyway.) The upshot is that **AC(S)** correctly discounts $R_{t_0} = 1$ as an actual cause of $A_{t_1} = 1$ in **Overlapping**.

The final point to note concerning **Overlapping** is that the definitions adopted here correctly count $U_{t_0} = 1$ as an actual cause of $A_{t_1} = 1$. To see this, consider the model $\mathcal{M}_{\textbf{Overlap}}$, described above and consider the partition of $\mathcal{V}_{\textbf{Overlap}} \backslash U_{t_0}, A_{t_1}$ such that $\mathbf{W} = \{A_{t_0}, R_{t_0}, R_{t_1}\}$ and $\mathbf{Z} = \{U_{t_1}\}$ and note that de facto probability-raising is revealed when we hold fixed the variables in **W** at their actual values. That is:

$$P(A_{t_1} = 1 | do(U_{t_0} = 1 \& A_{t_0} = 0 \& R_{t_0} = 1 \& R_{t_1} = 1)) > P(A_{t_1} = 1 | do(U_{t_0} = 0 \& A_{t_0} = 0 \& R_{t_0} = 1 \& R_{t_1} = 1)) \quad (21)$$

That is, in the actual circumstances in which the Ra-226 atom is present both at t_0 and t_1 and there's no alpha particle present at t_0, the probability of an alpha particle being present at t_1 is greater if the U-238 is present at t_0 than if it isn't.

Note that this de facto dependence isn't neutralized, because the following inequality also holds:

$$P(A_{t_1} = 1 | do(U_{t_0} = 1 \& U_{t_1} = 0 \& A_{t_0} = 0 \& R_{t_0} = 1 \& R_{t_1} = 1)) > P(A_{t_1} = 1 | do(U_{t_0} = 0 \& A_{t_0} = 0 \& R_{t_0} = 1 \& R_{t_1} = 1)) \quad (22)$$

That is, in the actual circumstances in which the Ra-226 atom is present both at t_0 and t_1 and there's no alpha particle present at t_0, the probability of an alpha particle being present at t_1 is greater if the U-238 atom is present at t_0 and absent at t_1 than if it isn't even present at t_0.

Definition **AC(M-R)** thus counts $U_{t_0} = 1$ as an actual cause of $A_{t_1} = 1$ relative to the model $\mathcal{M}_{\textbf{Overlap}}$. I've already noted that $\mathcal{M}_{\textbf{Overlap}}$ satisfies (Serious Possibilities), (Partition), (Independence), (Naturalness), and (Veridicality). Moreover, given that the actual value of U_{t_1} doesn't neutralize the de facto probability raising of $A_{t_1} = 1$ by $U_{t_0} = 1$ and given that, as Schaffer notes, U-238 decay doesn't occur via some intermediate process which could be incomplete, it doesn't seem that any variable or variables (representing reasonably natural states of affairs, etc.) could be added to $\mathcal{M}_{\textbf{Overlap}}$ that would neutralize this de facto probability-raising. (Nor does it seem that the existence of non-neutralized de facto probability-raising in $\mathcal{M}_{\textbf{Overlap}}$ rests upon not representing the events involved in a fine-grained enough way.) Consequently, (Stabilty) is also satisfied. **AC(S)** therefore correctly counts the t_0 presence of the U-238 as an actual cause *simpliciter* of the t_1 presence of the alpha particle.

7.4. Bogus Preemption

The final example that we'll consider here is one given by Hitchcock [11, p. 519] that is an example we might call 'bogus preemption'. Hitchcock offers this as an example that causes problems for the SEM approaches to deterministic actual causation presented in [10], [8], and [9]. But it also arguably poses a problem for the analysis of actual causation introduced above:

Bogus Preemption

"Assistant Bodyguard puts a harmless antidote into Victim's coffee. Buddy then poisons the coffee, using a type of poison that is normally lethal, but which is countered by the antidote. Buddy would not have poisoned the coffee if Assistant had not administered the antidote first. Victim drinks the coffee and survives."

Hitchcock [11, p. 520] reports that "[m]any people, but by no means all, have the intuition that Assistant's adding the antidote to the coffee is not a cause of Victim's survival."

Suppose that we wish to accommodate this intuition.[32] Then modification to **AC(M-R)** is needed. That's because the case is isomorphic to a preemption case. (Intuitively: just as my placing my U-232 near to the Geiger counter preempts your placing your Th-228 near to the Geiger counter – a non-actual event – as an actual cause of the bomb's exploding in **Probabilistic Preemption**, so Assistant's putting the antidote into the coffee preempts Buddy's not putting the poison in – a non-actual event – as a cause of Victim's survival. Moreover, a variable representing the poison's undergoing the nullifying reaction with the antidote plays a role analogous to the Geiger counter's reaching the threshold reading in **Probabilistic Preemption**.) The fact that the case is isomorphic to a preemption case means that the analysis given here, just like the SEM analyses of deterministic actual causation given in [10], [8], and [9], treats Assistant's action as an actual cause of Victim's survival.[33]

To see the trouble posed for the present analysis more clearly, consider a model $\mathcal{M}_{\textbf{BogusPre}}$ with a variable set $\mathcal{V}_{\textbf{BogusPre}}$ comprising the binary variables A, which takes the value $A = 1$ if Assistant puts the antidote in the coffee, and $A = 0$ otherwise; P, which takes the value $P = 1$ if Buddy puts the poison in the coffee, and $P = 0$ otherwise; N, which takes the value $N = 1$ if the poison undergoes the nullifying reaction with the antidote, and $N = 0$ otherwise; and S which takes the value $S = 1$ if Victim survives, and $S = 0$ otherwise. And suppose that the $do(\cdot)$ function associated with $\mathcal{M}_{\textbf{BogusPre}}$ is the verdical one relative to the world described in **Bogus Preemption**.

Now suppose that we evaluate the claim that $A = 1$ is an actual cause of $S = 1$ relative to $\mathcal{M}_{\textbf{BogusPre}}$. Consider the partition of $\mathcal{V}_{\textbf{BogusPre}} \setminus A, S$ such that $\mathbf{W} = \{P\}$ and $\mathbf{Z} = \{N\}$. And note that holding the variable in $\mathbf{W}$ fixed at its actual value, $P = 1$, reveals de facto probability-raising of $S = 1$ by $A = 1$:

---

[32] I'm not convinced that the payoff of doing so is worth the costs associated with the standard approach to accommodating this intuition that is taken within the SEM literature on deterministic actual causation. (And I don't have any novel approach of my own to propose.) Blanchard and Schaffer [49] describe these costs. The apparent equivocality of intuition about the case suggests that the payoff is not so great after all. Still, in what follows, I shall show that the present account is (once appropriately tweaked) as well-placed as its deterministic cousins to accommodate an intuition that Assistant's act was not an actual cause.

[33] In fact, Halpern and Pearl [9, Section 5] suggest a refinement of their analysis, which is roughly along the lines of that discussed below, which may allow them to avoid this result.

$$P(S = 1|do(A = 1\&P = 1)) > P(S = 1|do(A = 0\&P = 1)) \quad (23)$$

That is, holding fixed that Buddy puts poison in the coffee, the probability of Victim surviving is higher if Assistant has put the antidote in than if she hasn't. Moreover, this de facto probability-raising is not neutralized, because:

$$P(S = 1|do(A = 1\&N = 1\&P = 1)) > P(S = 1|do(A = 0\&P = 1)) \quad (24)$$

That is, holding fixed that Buddy puts poison in the coffee, the probability of Victim surviving is higher if Assistant puts the antidote in and the antidote nullifies the poison than if Assistant had never put the antidote into the coffee in the first place.

So **AC(M-R)** counts $A = 1$ as an actual cause of $S = 1$ relative to $\mathcal{M}_{\mathbf{BogusPre}}$. $\mathcal{M}_{\mathbf{BogusPre}}$ satisfies (Partition), (Independence), and (Naturalness). By stipulation, it also satisfies (Veridicality) for the scenario in question. It doesn't seem that there's any unmodelled neutralizer of the de facto probability-raising of $S = 1$ by $A = 1$ (nor does it seem that any fine-graining of the variables involved will reveal this de facto probability-raising to be spurious, as in the case of **Determinable Effect**), so it seems that $\mathcal{M}_{\mathbf{BogusPre}}$ satisfies (Stability). On the assumption that it satisfies (Serious Possibilities) (more on this in a moment), **AC(S)** counts $A = 1$ as an actual cause *simpliciter* of $S = 1$, which, to some at least, seems an implausible result.

Once it was noticed that the SEM analyses of deterministic actual causation given by Hitchcock [10] and Halpern and Pearl [8] also yield this result, advocates of the SEM approach attempted to deal with it by putting restrictions on what values other variables in a model could be held fixed at in evaluating whether the value of one variable is an actual cause of another.[34] In terms of the approach pursued here, the idea would be to say that unneutralized de facto probability-raising revealed by holding fixed the elements of $\mathbf{W}$ at their actual values $\mathbf{W} = \mathbf{w}^*$ is only indicative of actual causation if a certain further condition is satisfied. The idea would be to rule that, for the partition of $\mathcal{V}_{\mathbf{BogusPre}} \backslash A, S$ such that $\mathbf{W} = \{P\}$ and $\mathbf{Z} = \{N\}$, this further condition is not satisfied for the actual value, $P = 1$, of the variable in $\mathbf{W}$.

An idea developed by, among others, Halpern ([14], [57]) and Halpern and Hitchcock [16] is that this further condition can be framed in terms of considerations of *normality*. Applied to the present definitions, the idea is, when assessing whether $X = x$ rather than $X = x'$ is an actual cause of $Y = y$, (unneutralized) de facto dependence revealed by holding fixed the variables $\mathbf{W}$ at their actual values $\mathbf{W} = \mathbf{w}^*$ is indicative of actual causation iff the most normal world in which $X = x'$ and $\mathbf{W} = \mathbf{w}^*$ holds – the 'witness' world – is at least as normal as the actual world.[35] More specifically, the modified **AC(M-R)** would go as follows:

---

[34] Another case that has been used to motivate this sort of modification is identified by Hiddleston [56, p. 32] and dubbed 'bogus prevention' by Hitchcock [11, p. 523]. This is isomorphic to overdetermination cases, but only poses a problem to accounts that treat overdeterminers as causes, which the present account does not.

[35] Halpern and Hitchcock [17] adopt a more radical approach: they take actual causation to be gradable, with the degree of normality of the witness world determining the degree to which $X = x$ is an actual cause of $Y = y$ (when $Y = y$ counterfactually depends upon $X = x$ holding suitable factors fixed). The notion that actual causation might be gradable is a complication that I shall not explore here, though I don't see any reason to think that the proposal that the degree of normality of the witness determines the degree of actual causation wouldn't transfer smoothly to the present analysis.

AC(M-R)′

Where $x, x' \in \mathcal{R}(X)$ and $y \in \mathcal{R}(Y)$, $X = x$ rather than $X = x'$ is an actual cause *of $Y = y$ relative to a model $\mathcal{M} = \langle \mathcal{V}, \mathcal{R}, \mathcal{C}, \Omega, \mathcal{F}, do(\cdot) \rangle$ (with $X, Y \in \mathcal{V}$) in world θ iff there is a partition $(\mathbf{Z}, \mathbf{W})$ of $\mathcal{V} \backslash X, Y$ such that, where $\mathbf{W} = \boldsymbol{w}^*$ are the actual values the variables in $\mathbf{W}$ take in θ, the witness to $X = x' \& \mathbf{W} = \boldsymbol{w}^*$ is at least as normal as θ, and for all subsets $\mathbf{Z}'$ of $\mathbf{Z}$ (where $\mathbf{Z}' = \boldsymbol{z}^*$ are the actual values that the variables in $\mathbf{Z}'$ take in θ):*

(IN) $P(Y = y | do(X = x \& \mathbf{W} = \mathbf{w}^* \& \mathbf{Z}' = \mathbf{z}^*)) > P(Y = y | do(X = x' \& \mathbf{W} = \mathbf{w}^*))$

Halpern and Hitchcock suggest that various sorts of normality feed into determining the overall normality of a world, including statistical normality, default behaviors, moral and social norms, and norms of proper functioning.

As applied to the present example, the idea would be that a world in which Assistant doesn't put the antidote into Victim's coffee but Buddy nevertheless puts the poison in is an abnormal one, and more abnormal than the actual world in which Assistant puts the antidote in and Buddy puts the poison in. This might be justified on the grounds that Buddy's putting the poison in despite Assistant's not putting the antidote in would be a violation of a moral norm (or at least more of a violation of a moral norm than her putting the poison in when Assistant had already put the antidote in). Consequently, the fact that de facto probability-raising of $S = 1$ by $A = 1$ is revealed when it is held fixed that $P = 1$ is not indicative of $A = 1$ being an actual cause of $S = 1$.

In order for this proposal to work it would have to preserve the desirable results that we get in other cases. Thus in **Probabilistic Preemption**, for instance, it would be important that the de facto probability-raising of the explosion by my placing my U-232 near to the Geiger counter that is revealed by holding fixed your non-placement of your Th-228 near to the Geiger counter still be counted as genuinely indicative of actual causation. To get this result it would be necessary that, according to the appropriate notion of normality, a world in which you don't place your Th-228 near to the Geiger counter and I don't place my U-232 near to the Geiger counter either is at least as normal as the actual world, in which I place my U-232 but you don't place your Th-228 near to the Geiger counter. But this is so on quite a reasonable sense of normality: after all, your placing your Th-228 near to the Geiger counter violates moral norms (given the setup involving the bomb) and also, simply in virtue of being a positive action rather than an inaction, is a non-default behavior.

Two obvious questions present themselves at this point. First, does this appeal to 'normality' do any work that the appeal to 'serious possibilities' (in defining model appropriateness) does not already do? Second, if it does, can we drop the appeal to serious possibilities in favour of this appeal to normality? I'd give a cautious affirmative answer to each of these questions (cautious because this partly depends upon how the account of 'normality' is worked out and what counts as a 'serious possibility').

Regarding the first question, one might ask, for instance, whether $\mathcal{M}_{\mathbf{BogusPre}}$ violates (Serious Possibilities) so that we can conclude that the unneutralized de facto probability-raising of $S = 1$ by $A = 1$ in this model fails to be indicative of actual causation. If so, we wouldn't need to appeal to the normality of the witness world to secure the latter result. Perhaps, for instance, we could argue that a model including both A and P is inappropriate model because such a model allows the representation of a non-serious possibility: namely, the possibility that Buddy poisons the coffee even though Assistant hasn't put the antidote in.

But I don't think that this strategy of trying to secure the desired result by simply appealing to (Serious Possibilities) will work. (Serious Possibilities) was never intended by its advocates to rule as inappropriate models that include only variables that represent salient parts of the causal structure being modelled, as both S and A do for **Bogus Preemption**. Rather, it was introduced to rule as inappropriate models including variables representing highly non-salient factors, such as whether or not McCartney intercepted the bullets in the case of the JFK assassination. Indeed, interpreting (Serious Possibilities) as ruling out models including the former sort of variable and not just those including the latter sort would fatally undermine treatments of actual causation that appeal to SEMs or PCMs. That's because even models that seem like perfectly reasonable representations of just the salient events in causal scenarios are typically such that there is *some* assignment of values to the variables they include that corresponds to a rather remote possibility.

Consider, for instance, $\mathcal{M}_{\mathbf{ProbPre}}$. And note that there seems to be no reason why we should suppose that your commitment to not place your Th-228 near to the Geiger counter if I place my U-232 near to the Geiger counter in **Probabilistic Preemption** should be any weaker than Buddy's commitment not to poison the coffee if Assistant doesn't put the antidote in in **Bogus Preemption**. But, by including M and Y in $\mathcal{M}_{\mathbf{ProbPre}}$, we allow the representation of the possibility that I place my U-232 near the Geiger counter *and* you place your Th-228 near to the Geiger counter. Ruling such a model to be non-apt because it violates (Serious Possibilities) would be highly problematic, since it's only by including both M and Y that we can represent the de facto probability-raising of $E = 1$ by $M = 1$ that is indicative of actual causation in this case. As we've seen, interpreting (Serious Possibilities) so that it doesn't rule out a model including M and Y, but introducing a 'naturalness' restriction on what settings of $\mathbf{W}$ are such that unneutralized de facto dependence relative to them is indicative of actual causation plausibly preserves the correct result in this case.

On the other hand, once the appeal to naturalness has been invoked, it's plausible that we can dispense with (Serious Possibilities). If we do so, a model of the JFK assassination counts as appropriate even if it includes a variable corresponding to whether or not McCartney got in the way of the bullets. Still, the appeal to naturalness plausibly rules out McCartney's failure to intercept the bullets from counting as an actual cause of JFK's death because a world in which McCartney intercepts them would presumably be an abnormal one by any reasonable standards.

8. Conclusion

In this paper, PCMs have been drawn upon in providing an analysis of actual causation. The aspiration has been to provide a treatment of probabilistic actual causation that matches the sophistication and success of those treatments of deterministic actual causation that draw upon SEMs, whilst also paying due attention to the unique challenges that arise in the probabilistic case. It has been shown that this analysis gives the correct treatment of a range of test cases, some of which have been raised in the literature as posing problems for traditional accounts of probabilistic actual causation, and some of which suggest themselves as challenging cases because they are analogues of problem cases widely discussed in the deterministic actual causation literature.

Funding: This research received no external funding.

Acknowledgments: For comments on earlier versions of this material, I would like to thank audiences at the Universities of Edinburgh, Hamburg, and Bristol, at the Metaphysics of Entanglement Speaker Series at Oxford, the British Society for the Philosophy of Science Speaker Series, and the Choice Group at London School of Economics.

Conflicts of Interest: The author declares no conflict of interest.

1. Eagle, A. Generic Causation. Unpublished Manuscript; URL: ⟨http://philpapers.org/rec/EAGGC⟩.
2. Mackie, J. Causes and Conditions. *American Philosophical Quarterly* **1965**, *2*, 245–64.
3. Lewis, D. Causation. *Journal of Philosophy* **1973**, *70*, 556–67.
4. Salmon, W. *Scientific Explanation and the Causal Structure of the World*; Princeton University Press: Princeton, NJ, 1984.
5. Menzies, P. Probabilistic Causation and Causal Processes: A Critique of Lewis. *Philosophy of Science* **1989**, *56*, 642–663.
6. Pearl, J. *Causality: Models, Reasoning, and Inference*, first ed.; Cambridge University Press: Cambridge, 2000.

7. Pearl, J. *Causality: Models, Reasoning, and Inference*, second ed.; Cambridge University Press: Cambridge, 2009.
8. Halpern, J.Y.; Pearl, J. Causes and Explanations: A Structural-Model Approach. Part I: Causes. Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence; Morgan Kaufmann: San Francisco, CA, 2001; pp. 194–202.
9. Halpern, J.Y.; Pearl, J. Causes and Explanations: A Structural-Model Approach. Part I: Causes. *British Journal for the Philosophy of Science* **2005**, *56*, 843–87.
10. Hitchcock, C. The Intransitivity of Causation Revealed in Equations and Graphs. *Journal of Philosophy* **2001**, *98*, 194–202.
11. Hitchcock, C. Prevention, Preemption, and the Principle of Sufficient Reason. *Philosophical Review* **2007**, *116*, 495–532.
12. Woodward, J. *Making Things Happen: A Theory of Causal Explanation*; OUP: Oxford, 2005.
13. Glymour, C.; Wimberly, F. Actual Causes and Thought Experiments. In *Causation and Explanation*; Campbell, J.K.; O'Rourke, M.; Silverstein, H., Eds.; MIT Press: Cambridge, MA, 2007; pp. 43–67.
14. Halpern, J.Y. Defaults and Normality in Causal Structures. Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning; Brewka, G.; Lang, J., Eds.; AAAI Press: Menlo Park, CA, 2008; pp. 198–208.
15. Halpern, J.Y. Appropriate Causal Models and Stability of Causation. Proceedings of the Fourteenth International Conference on Principles of Knowledge Representation and Reasoning; AAAI Press: Palo Alto, CA, 2014; pp. 198–207.
16. Halpern, J.Y.; Hitchcock, C. Actual Causation and the Art of Modeling. In *Heuristics, Probability and Causality*; Dechter, R.; Geffner, H.; Halpern, J.Y., Eds.; College Publications: London, 2010; pp. 383–406.
17. Halpern, J.Y.; Hitchcock, C. Graded Causation and Defaults. *British Journal for the Philosophy of Science* **2015**, *66*, 413–457.
18. Weslake, B. A Partial Theory of Actual Causation **2016**. Forthcoming in *British Journal for the Philosophy of Science*.
19. Good, I.J. A Causal Calculus (I). *British Journal for the Philosophy of Science* **1961**, *11*, 305–18.
20. Good, I.J. A Causal Calculus (II). *British Journal for the Philosophy of Science* **1961**, *12*, 43–51.
21. Reichenbach, H. *The Direction of Time*; Dover: Mineola, NY, 1971.
22. Suppes, P. *A Probabilistic Theory of Causality, Acta Philosophica Fennica*; North-Holland: Amsterdam, 1970.
23. Lewis, D. *Philosophical Papers*; Vol. 2, Oxford University Press: Oxford, 1986.
24. Menzies, P. Probabilistic Causation and the Pre-emption Problem. *Mind* **1996**, *105*, 85–117.
25. Hitchcock, C. Do All and Only Causes Raise the Probabilities of Effects? In *Causation and Counterfactuals*; Collins, J.; Hall, N.; Paul, L., Eds.; MIT Press: Cambridge, MA, 2004; pp. 403–417.
26. Goldszmidt, M.; Pearl, J. Rank-Based Systems. Proceedings of the Third International Conference on Knowledge Representation and Reasoning; Morgan Kaufmann: San Mateo, CA, 1992; pp. 661–672.
27. Lewis, D. Counterfactual Dependence and Time's Arrow. *Noûs* **1979**, *13*, 455–76.
28. Glynn, L. Of Miracles and Interventions. *Erkenntnis* **2013**, *78*, 43–64.
29. Hitchcock, C. Farewell to Binary Causation. *Canadian Journal of Philosophy* **1996**, *26*, 267–82.
30. Hitchcock, C. The Role of Contrast in Causal and Explanatory Claims. *Synthese* **1996**, *107*, 395–419.
31. Schaffer, J. Contrastive Causation. *Philosophical Review* **2005**, *114*, 327–58.
32. Schaffer, J. Causal Contextualisms. In *Contrastivism in Philosophy*; Blaauw, M., Ed.; Routledge: New York, 2013; pp. 35–63.
33. Schaffer, J. Causes as Probability Raisers of Processes. *Journal of Philosophy* **2001**, *98*, 75–92.
34. Anscombe, E. *Causality and Determination*; CUP: Cambridge, 1971.
35. Feynman, R. *The Character of Physical Law*; BBC: London, 1965.
36. Lewis, D. Postscripts to "Causation". In *Philosophical Papers*; Lewis, D., Ed.; Oxford University Press: Oxford, 1986; Vol. 2, pp. 172–213.
37. Loewer, B. Determinism and Chance. *Studies in History and Philosophy of Science Part B* **2001**, *32*, 609–620.
38. Hoefer, C. The Third Way on Objective Probability: A Sceptic's Guide to Objective Chance. *Mind* **2007**, *116*, 549–96.
39. Ismael, J. Probability in Deterministic Physics. *Journal of Philosophy* **2009**, *106*, 89–108.
40. Ismael, J. A Modest Proposal About Chance. *Journal of Philosophy* **2012**, *108*, 416–42.

41. Glynn, L. Deterministic Chance. *The British Journal for the Philosophy of Science* **2010**, *61*, 51–80.
42. Emery, N. Chance, Possibility, and Explanation. *British Journal for the Philosophy of Science* **2015**, *66*, 95–120.
43. Frigg, R.; Hoefer, C. The Best Humean System for Statistical Mechanics. *Erkenntnis* **2015**, *80*, 551–574.
44. Hitchcock, C. A Tale of Two Effects. *Philosophical Review* **2001**, *110*, 361–96.
45. Kvart, I. Causation: Probabilistic and Counterfactual Analyses. In *Causation and Counterfactuals*; MIT Press: Cambridge, MA, 2004; pp. 359–386.
46. Glynn, L. A Probabilistic Analysis of Causation. *British Journal for the Philosophy of Science* **2011**, *62*, 343–392.
47. Yablo, S. De Facto Dependence. *Journal of Philosophy* **2002**, *99*, 130–148.
48. Yablo, S. Advertisement for a Sketch of an Outline of a Prototheory of Causation. In *Causation and Counterfactuals*; MIT Press: Cambridge, MA, 2004; pp. 119–137.
49. Blanchard, T.; Schaffer, J. Cause without Default. In *Making a Difference*; Beebee, H.; Hitchcock, C.; Price, H., Eds.; OUP: Oxford, 2016; pp. 175–214.
50. List, C.; Pivato, M. Emergent Chance. *Philosophical Review* **2015**, *124*, 119–152.
51. Halpern, J.Y. A Modification of the Halpern-Pearl Definition of Causality. Proceedings of the 24^{th} International Joint Conference on Artificial Intelligence; Yang, Q.; Wooldridge, M., Eds.; AAAI Press: Menlo Park, CA, 2015; pp. 3033–3033.
52. Schaffer, J. Trumping Preemption. *Journal of Philosophy* **2000**, *97*, 165–181.
53. Lewis, D. Causation as Influence. In *Causation and Counterfactuals*; Collins, J.; Hall, N.; Paul, L., Eds.; MIT Press: Cambridge, MA, 2004; pp. 75–106.
54. Bigaj, T. Causation Without Influence. *Erkenntnis* **2012**, *76*, 1–22.
55. Schaffer, J. Overlappings: Probability-Raising Without Causation. *Australasian Journal of Philosophy* **2000**, *78*, 40–6.
56. Hiddleston, E. Causal Powers. *British Journal for the Philosophy of Science* **2005**, *56*, 27–59.
57. Halpern, J.Y. *Probabilistic Causality*; MIT Press: Cambridge, MA, 2016.