Active-path token causation extended to probabilistic cases

Charles R. Twardy, Kevin B. Korb*
Graham Oppy & Toby Handfield†
Monash University, Australia

December 8, 2004

Abstract

We present a probabilistic extension to active-path analyses of token causation (Halpern and Pearl, 2001, 2004; Hitchcock, 2001). We also make the previous accounts more explicit, presenting them in computational form. The extension uses the generalized notion of intervention presented in (Korb, Hope, Nicholson, and Axnick, 2004): we allow an intervention to set any probability distribution over the possible values, not just a single value. The resulting account can handle a wide range of examples. We do not claim the account is complete — only that it fills an obvious gap in previous active-path approaches. It still succumbs to recent counterexamples by Hiddleston (forthcoming 2004), because it does not explicitly consider causal processes. We claim three benefits: a detailed comparison of three active-path approaches, a probabilistic extension for each, and an explicit computational form.

Causal models represent type causation, such as the general effect of smoking on lung cancer. Recently, Halpern & Pearl (2001; 2004) and Hitchcock (2001) have shown how they can also be used to analyze token causation. The accounts make sense of token cases where things happen "the hard way", and naturally handle troublesome cases involving pre-emption and multiple paths. Hitchcock's account is more straightforward, but slightly less general. Unfortunately, both accounts assume deterministic networks, which is unnecessary and contrary to the spirit of causal models.

The basic trick for token causation is to make use of a causal model's natural type-causation semantics: C is a cause iff there is a context in which a change in the value of C would make a difference to E. The problem is that in token causation, we must restrict which contexts we can consider. Whether a token event c causes e depends importantly upon the actual context in which c and e occur. The accounts we present, and therefore our probabilistic extension, are just attempts to set out these restrictions. On the accounts presented here, the relevant contexts are those in the **redundancy range** for variables not on the **active path**.

We acknowledge recent counterexamples (Hiddleston, forthcoming 2004) showing that redundancy ranges sometimes pick the wrong context. It is likely that a complete account must make reference to causal processes, and we develop one in a companion piece (Korb et al., 2005). Nevertheless, this account is valuable

*Computer Science & Software Engineering, email:{ctwardy,korb}@csse.monash.edu.au
†Philosophy, email:{Graham.Oppy,Toby.Handfield}@arts.monash.edu.au

for greatly extending active-path approaches. By presenting algorithms for the theory, it also opens the way for automatically comparing accounts of causation on large sets of examples, providing philosophers with much better metrics for comparison than the usual handful of successes or failures presented in any paper.

1 Deterministic causation

We begin with an example from Hitchcock (2001). A boulder falls F towards a hiker. The hiker sees the boulder in time, and ducks D, therefore surviving, S. We represent the causal variables as nodes in a graph, and the causal links as arrows between them. We specify the exact dependencies with a series of simple structural equations. So:

$$\begin{tikzcd} F \arrow[rr] \arrow[dr] & & S \\ & D \arrow[ur] & \end{tikzcd}$$

$$\begin{aligned} F &= 1 \\ D &= F \\ S &= \neg F \vee D \end{aligned}$$

The first equation gives an initial condition: in the actual world, the boulder falls. The second says that the hiker ducks iff the boulder falls. The third says the hiker survives if he ducks, or if the boulder does not fall. The equations are asymmetric because '=' stands for "is determined by", not mere algebraic equality.

Graph plus equations plus possible values of the root variables specify our causal model. The model encompasses all possibilities we are willing to entertain seriously. One way to think about the model is a mathematically precise set of possible worlds. The model defines the set of counterfactuals which can be entertained, and provides truth conditions for evaluating them.

Note: For simplicity we presume that all descendants of E are excluded from the model, or at any rate, never observed or otherwise fixed. We do this to avoid confusion. Knowing descendants of E will tell us what happened, and obscure the causal impact of C. We would get the right answers to the wrong questions.

1.1 Simple Token Causation (H1)

In the Boulder example, we might ask whether ducking causes the hiker to survive. Intuitively, the answer is "yes": had the hiker not ducked, he would not have survived. However, a naïve account might allow counterfactuals to "backtrack" like so: had the hiker not ducked, the boulder would not have fallen, so ducking made no difference to survival. David Lewis ruled out backtracking in his original counterfactual analysis of causation (1973), and to our knowledge, no one has held such an account.

Almost any non-backtracking account says that ducking saved the hiker, and the semantics of causal models preclude backtracking. We merely replace the

equation for D with a particular value—holding fixed the background variables (F) at their actual values—, and let the remaining machinery calculate the answer. In this case, F is unaltered, as is the equation for S. Then, $S = 1$ when $D = 1$ and $S = 0$ when $D = 0$, so ducking makes a difference. Therefore, ducking caused survival.

When we replace an equation like this, we say we have **blocked** the path to D, because there is no longer any counterfactual dependence of D on F. In effect, by replacing D with a particular value ($D = 0$), we have removed the arc $F \longrightarrow D$, or at least blocked the influence travelling along it. Such "arc surgery" gives an intuitive way to entertain counterfactuals about causal relationships. Given our causal model, the closest possible world where the hiker did not duck is just the same model with $D = F$ (an arc) replaced by $D = 0$ (a fixed value). The rest of the causal structure is left intact.

So far this method of counterfactual reasoning is in agreement with Lewisian counterfactual semantics. Next, we might ask whether the boulder's fall caused the hiker to survive. Here, intuition says "no", and indeed there is no counterfactual dependence. However, in order to handle several counterexamples, Lewis (1973) abandoned simple counterfactual dependence in favor of *chains* of dependence. These "Lewis chains" entail that causation is transitive. Therefore, because S depends on D and D depends on F, Lewis would say that the boulder's fall caused the hiker to survive.

Causal model counterfactuals have all the advantages of Lewis chains without entailing transitivity. Let us see how Hitchcock (2001) handles the case. We define an Active Path Token Cause.

Definition 1 (APT-C:) *$C = c$ was a token cause of $E = e$ iff there was an* ***active path*** *from C to E.*

A **path** is a walk from C to E along the arrows, in the direction of the arrows. A path is **active** if it was making a difference. However, we will have to distinguish between *strongly* and *weakly* active paths. We leave weakly active paths for the next section.

A path is *strongly* active if $C = c$ makes a difference to $E = e$ when all other paths between C and E are blocked by setting variables along those paths to their actual values. "Making a difference" has the usual counterfactual meaning: Had C been different, E would have been different, for at least *some* $c' \neq c$.

In Boulder, neither of the paths from F to S is strongly active. (It also turns out that they are not weakly active.) First, we hold D fixed at its actual value ($D = 1$), and note that the direct path $F \longrightarrow S$ is inactive: given that the hiker ducked, the boulder has no effect. Then, we fix all paths except FDS. There being no variables to fix along the other path, we fix nothing, and note that overall, S does not depend on F. We happily conclude that F did not cause S. "Intuitively, the falling boulder does not save Hiker's life because without it, Hiker's life would not have been endangered in the first place" (Hitchcock, 2001, page ??).

In general, paths can be long chains like $C \longrightarrow X \longrightarrow Y \longrightarrow E$. We want to leave exactly one path open, such as the path through D, and block all others. A straightforward (if inefficient) way to do this is to fix *all* ancestors of E, except those along ϕ. This is the **background** for E, which we shall call β. (It is common to use π for the *parents* of E, so if we use Π for all the ancestors, then $\beta = \Pi \backslash \phi$.)

Let c and c' be states of C, with c the actual state. Let e be the actual state of E. We use the notation I_c to mean that we intervene to set $C = c$, replacing the equation for C with the value c. Likewise, $I_{c'}$ means we intervene to set $C = c'$. Etc.

Then we define:

Definition 2 (Strongly active path:) *a path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$ is strongly active in causal model $\mathcal{M}$ iff I_c makes a difference to E (relative to $I_{c'}$) in the β-actual model, where the β-actual model replaces all the equations for the* ***background*** *of E with their actual values.*

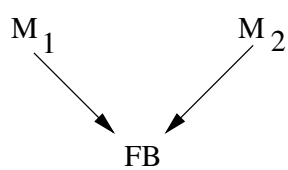
In Hitchcock's examples, all paths to E start at C. But in general, there can also be external causes. The simplest case is $C \longrightarrow E \longleftarrow D$. Our definition rightly requires us to block *all* other paths to E, including those originating from other causes.

If we limit active paths to *strongly* active paths, APT-C gives us Hitchcock's simple account of causation. Following Hiddleston (forthcoming 2004), we shall call this **H1**.

Unfortunately, this simple account cannot handle symmetric overdetermination nor late cutting. We must invoke weakly active paths.

1.2 Symmetric Overdetermination (H2)

Symmetric overdetermination happens when there are multiple sufficient causes. One such case is Matches. Two lit matches (M_1 and M_2) are dropped in the forest. Either one would suffice to burn it down. The forest burns down (FB).



$$\begin{aligned} M_1 &= 1 \\ M_2 &= 1 \\ FB &= M_1 \vee ML_2 \end{aligned}$$

Did M_1 cause the forest to burn down? Given M_2, it made no difference, so a naïve counterfactual account says no. Hitchcock's simple account also says no. Using Definition 2, we fix M_2 at its actual value, and find that the forest burns down regardless of how we set M_1. The problem being symmetric, M_2 likewise makes no difference. We have a problem.

Lewis set aside these completely symmetric cases claiming he had no clear intuitions (1973, ??), and it was "spoils to the victor". Not only does our analysis differ from his, but we *do* have clear intuitions. In fact, both matches contributed, even though neither alone was necessary. It is only by chance that the actual circumstances have masked the probabilistic dependency. The causal structure shows how to reveal it.

Note that we can let M_2 be 0 without affecting the result—the forest still burns down. A variable often has several values which make no difference on their own. Hitchcock calls this set the **redundancy range**. Let us set $M_2 = 0$. We immediately see the hidden dependency: now M_1 makes a difference! That is what it means to be a token cause: to actually contribute to the causal history, even if your effect is accidentally masked. That is another reason why we cannot identify token causation with actual probabilistic relevance.

Noting that an instantiated model is one with actual values for its variables, we define:

Definition 3 (Redundancy range:) *Given an instantiated causal model $\mathcal{M}$, a path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$, and the background $\beta = \Pi \backslash \phi = \{B_1, \ldots, B_m\}$, the set of values $\mathbf{b}' = \{b_1, \ldots, b_m\}$ lies in β's* redundancy range *for ϕ iff $I_{\mathbf{b}'}$ leaves ϕ unchanged.*

The redundancy range for path ϕ is the set of all possible ways of fixing the other variables so that they make no difference to anything in ϕ. Note that the actual values of β are also in the redundancy range!

Hitchcock defines a *weakly* active path to be one which is revealed by setting variables outside that path to values in their redundancy range. However, since the redundancy range includes the actual values, strongly active paths are (somewhat counterintuitively) also weakly active.

Here is Hitchcock's definition, rewritten in our terms.

Definition 4 (Weakly active path:) *a path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$ is weakly active in causal model $\mathcal{M}$ iff I_c makes a difference (relative to $I_{c'}$) to E in any β-redundant model, where a β-redundant model replaces all $B \in \beta$ with values in their redundancy range for instantiated ϕ. (Replacement is an intervention.)*

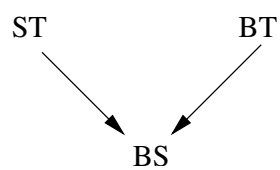
Using weakly active paths Hitchcock's definition can now handle cases of symmetric overdetermination: neither path in Matches is strongly active, but both are weakly active, so both matches are causes according to Definition 1. Hiddleston calls this account **H2**.

1.3 Late Cutting

Late cutting is a kind of pre-emption by overdetermination. It occurs when one of the potential causes pre-empts the other, precisely by completing first. A common example is Bottle:

> Suzy and Billy throw rocks at a bottle. Suzy's rock arrives first, breaking the bottle before Billy's rock, which would also have broken it.

We want to say that Suzy's throw caused the bottle to break, but that Billy's did not. First, note that the obvious model does not work.



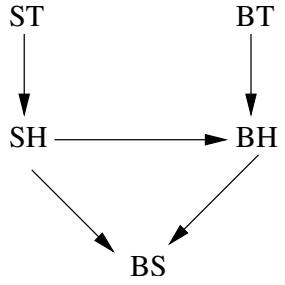
$$\begin{aligned} ST &= 1 \\ BT &= 1 \\ BS &= ST \vee BT \end{aligned}$$

This is the same model as Disjunctive Matches, so it gives the same answer: both Billy's throw and Suzy's throw are weakly active, so both caused the bottle to smash. But our intuitions *strongly* favor Suzy's throw over Billy's. Billy's throw was pre-empted.

Halpern and Pearl (2001) argue that we have misrepresented the case. According to them, since Matches is symmetric and Bottle is not, we cannot represent both with the same model. They argue that a proper model should

structurally show the key asymmetry in the example: Suzy's throw hitting prevents Billy's from hitting. (We argue elsewhere Korb et al. (2005) that we can keep this model if we attend to processes. Here we strive merely to give a probabilistic extension of active-path accounts, so we keep to their formulation.)

Halpern & Pearl suggest the following model:



$$\begin{aligned} ST &= 1 \\ BT &= 1 \\ SH &= ST \\ BH &= BT \wedge \neg SH \\ BS &= SH \vee BH \end{aligned}$$

This seems to be a good model, and it works. Suzy's path is strongly active: given that Billy's rock does not hit, had Suzy not thrown, the bottle would not have shattered. But Billy's path is not active.

It is not strongly active because in the actual case he does not hit the bottle. It is not weakly active because the redundancy range is degenerate.[1] The only path from BT to BS is $BT \longrightarrow BH \longrightarrow BS$. Everything else is background. If Suzy had not hit ($SH = 0$), then Billy would have ($BH = 1$). But then we have changed BH, which is on our path, so ($SH = 0$) is not in the redundancy range for ϕ. Given our interventions on SH, ST is irrelevant. The redundancy range is degenerate, so BT is not weakly active.

Therefore $BT \longrightarrow BH \longrightarrow BS$ was inactive. There being no other path from BT, Billy's throw did not cause the bottle to smash.

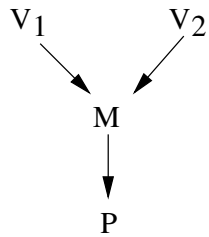
H2 can handle most cases in the philosophy literature. However, Halpern and Pearl (2004) present one which **H2** cannot handle.[2]

1.4 Problem and diagnosis

The Voting Machine example (Halpern and Pearl, 2004, Example A.3) shows that **H2** is too strict about the path ϕ. Suppose we model a simple voting scenario. Two people (V_1 and V_2) vote, and the measure passes (P) if at least one votes in favor. Both vote in favor, and the measure passes. This scenario is similar to Matches, and both votes cause the measure to pass. Suppose we now introduce a voting machine (M) that tabulates the votes first. Then our model is:

[1] Perhaps it is only a dundancy range?

[2] Our thanks to Chris Hitchcock for directing our attention here and saving us from trying to prove equivalence between **H2** and Halpern & Pearl.



$$\begin{aligned} V_1 &= 1 \\ V_2 &= 1 \\ M &= V_1 + V_2 \\ P &= M \geq 1 \end{aligned}$$

However, **H2** now says that neither $V_1 = 1$ nor $V_2 = 1$ causes $P = 1$. Like Matches, neither path is strongly active. Unlike Matches, they are not even weakly active, because the redundancy range is degenerate. The problem is V_1 and V_2 can affect P only via M, which is on the path. **H2** explicitly prevents the background variable from changing the value of variables along the path. This is too strong. As Halpern & Pearl note, "we cannot insist on the variables in [the path] remaining constant; instead, we require merely that changes in [the background] not affect [E]." We now present their account. We shall call it **H3**.

1.5 H3: Halpern & Pearl

Where Hitchcock forbids any alteration to the path, Halpern & Pearl allow any alteration which does not affect E, but then add a "resetting" clause that captures the crucial path asymmetry in cases of late cutting.

Here is our rendition of Halpern & Pearl's account.

Definition 5 (H3) *$C = c$ is a token cause of $E = e$ iff:*

AC1 *Actually, c and e.*

AC2 *There is a partition of variables into ϕ (path) and β (background), with $C \in \phi$, and some values c', b' such that:*

(A) *If we were to set c' and b', we would get $e' \neq e$.*

(B) *If we were to set c and b', we would still get e,* even if *we reset any of the other variables in ϕ to their actual values.*

This rendition is identical in effect, but with some minor changes for clarity.[3]

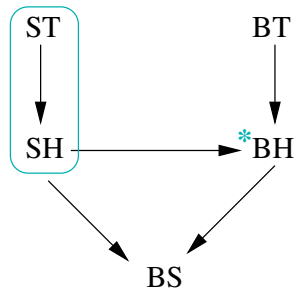
AC2(a) amounts to saying that there is a background state (b') where c can make a difference to e. Condition AC2(b) ensures that the background b' alone is insufficient to change e to e': it means that b' is in the redundancy range for e. The "even if" bit about resetting variables in ϕ to their actual values is only necessary in cases of late cutting.

[3] First, we omit the exogenous error terms, which serve mostly to complicate the expressions and assuage deterministic leanings. Second, we use helpful variable names. Since their set Z will only include variables between C and E, we call it ϕ for "path". Their set W will include all the ancestors of E not in Z, hence β for "background". Knowing their interpretation greatly aids the search for partitions! Finally, we presume C is a single variable, so do not need their minimality condition AC3.

Note: The "Background" β often includes variables *downstream* from C. That means our causal background to E — chosen to block alternate paths of influence from C to E — will likely include the *temporal* foreground of C.

In the augmented bottle-smashing model, Suzy's throw $ST = 1$ is a token cause of $BS = 1$ in the same way it was for Hitchcock's weakly-active paths: we use a redundancy range to reveal the dependency. Choosing $\{BT, BH\}$ as our background, and setting $BH = 0$, we see that ST makes a difference.

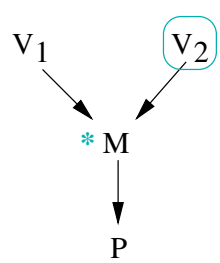
Billy's throw would similarly make a difference, save for the "resetting" clause. Let the background be $\{ST, SH\}$ (shown by the rounded box, below). Setting $SH = 0$, Billy's throw makes a difference.



However, the asterisk denotes we reset BH to its original value. Then $BS = 0$, in violation of AC2(B). Therefore this path is not an active path, and BT is not an actual cause.

When detailing Hitchcock's account, we took the expedient of fixing the *entire* background. That would still work here, H3 only requires that fixing *some* subset $\mathcal{B}$ of B allows c to make a difference. In the deterministic case, it doesn't matter: if fixing some $\mathcal{B}$ works, then there will also be a way to fix all of B. In the probabilistic case, that won't be true.

Before moving on to probabilities, we present the voting machine case in detail. In the hope that it will help readers working through the example, we present the solution as two opposing players taking turns in a short game. The "advocate" tries to show that some $C = c$ is a cause of $E = e$.[4] To this end, he nominates an active path, a value $c' \neq c$, and context b' such that given b', c yields e and c' yields e'. Next the "naysayer" tries to find a way of resetting the intermediate variables along that path to make it inactive (meaning that c no longer yields e).



ADV: $(V_1 = 1)$ $[c]$ is a cause of $(P = 1)$ $[e]$. The path is obviously $\phi = V_1 \longrightarrow M \longrightarrow P$. Background: $b' = V_2$, setting $V_2 = 0$. The cause makes a difference: c gives e, and $(V_1 = 0)$ $[c']$ gives $(P = 0)$ $[e']$.

NAY: Your b' has made $M = 1$. I reset it to its original value, $M = 2$.

ADV: No worries. $M = 2$ still gives $P = 1$, the original e.

[4] After all, every advocate must have a cause.

NAY: I have no other intermediate variables to try. You win.

Of course, this is still limited to deterministic causation. However, the probabilistic extension is straightforward.

2 HP: A probabilistic extension

The counterfactuals in **H1** through **H3** are of the form, "if not $(C = c)$, then not $(E = e)$." That will not work for probabilistic causation. In the Matches example, if each match acts independently with probability p of burning down the forest, then both matches have a greater probability of burning down the forest, but it still might not burn down. Both paths are active because they change the probability that the forest will burn down.

HP replaces that absolute counterfactual with a change in probabilities: "if not $(C = c)$, then $\Pr(E = e)$ would have been different.". We keep the same desideratum APT-C: c was a token cause (now meaning token causally *relevant*) iff there was an active path from c to e. We will, of course, redefine active paths to mean probabilistic differences.

We could call this "Probabilistic Token Causal Relevance" to distinguish it from the absolute version. But that would suggest there was a place for the non-probabilistic account, whereas the probabilistic version entirely subsumes it.

Following Hitchcock, we will first present the simple version. This simple version does not invoke weakly-active paths, so it still has trouble with some cases of symmetric overdetermination (like the original Matches example) and late cutting (like Bottle).

Parallelling Hiddleston, we will call these **HP1**, **HP2**, and **HP3**.

2.1 HP1: Strongly active paths

We keep the earlier definition of strongly active path. However, "makes a difference" now means that there is a change in probabilities, so it is worth rewriting. Define

$$\Delta_1 = \Pr(e|I_c) - \Pr(e|I_{c'})$$

and let $0 \leq \epsilon < 1$ denote some tolerance threshold below which we consider the change in probability to be meaningless. Then:

Definition 6 (Strongly active:) *a path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$ is strongly active in causal model $\mathcal{M}$ iff abs$(\Delta_1) > \epsilon$ in the β-actual model, where the β-actual model replaces all the equations for the* ***background*** *with their actual values.*

Recall that replacing the background is an intervention, so the equation for Δ_1 implicitly conditions on $I_\beta = I_{b_1,\ldots,b_m}$ Like before, this definition requires an explicit contrast class c'.[5] However, there is a difference at E. Our definition Δ is relative to the probability of the actual state e. So if c vs c' reverses the probabilities of e_1 and e_2, but leaves unchanged the probability of actual state e_3, we will count this as no effect.

While that seems more in keeping with the philosophy literature, one could also just measure the distance between the entire distribution $\Pr(E|I_c) - \Pr(E|I_{c'})$, using say Kullback-Leibler "distance" or Bhattacharyaa distance. For now we

[5] Because we have *intervened* to set β, we can let c' default to the "prior" distribution on C after fixing β. If we merely observed β, we might have "unblocked" probability flow from separate causes of E that share the same path.

suggest only that the most straightforward extension makes reference to the actual state e.

Roughly then, **HP1** says:

Definition 7 (HP1:) *Relative to c', c was token causally relevant to e iff for some path ϕ,* $\mathrm{abs}(\Delta_1) > \epsilon$.

The extra details in Definition 6 specify how we calculate Δ_1.

Speakers often want to know more than just that smoking is causally relevant. They want to know the direction (or valence) of the influence. However, C may promote E along some paths, and hinder E along other paths. Therefore in C may both promote and prevent E. (Of course, in any particular context the *net* effect will be either positive, negative, or too small to count.)

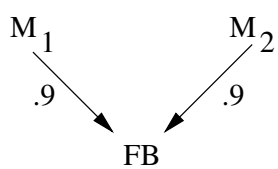
Definition 8 (Token Promotion) *c token promotes e iff for some path ϕ:*

1. *c is token causally relevant to e along ϕ, with*
2. *$\Delta_1 > 0$: (Relative to c', c increases the chance of e.)*

Definition 9 (Token Hindrance) *c token hinders e iff for some path ϕ:*

1. *c is token causally relevant to e along ϕ, with*
2. *$\Delta_1 < 0$: (Relative to c', c decreases the chance of e.)*

Consider a simple probabilistic version of Matches where M_1 alone results in $FB = 1$ with probability s, and M_2 alone with chance t. Then $FB = 0$ with chances $\sigma = 1 - s$ and $\tau = 1 - t$. Assume the forest cannot burn down spontaneously. Then if $s = t = .9$, and the matches act independently, our model might be:



$$\begin{aligned} M_1 &= 1 \\ M_2 &= 1 \\ \Pr(FB) &= 1 - \sigma^{M_1}\tau^{M_2} \end{aligned}$$

where the exponents just switch the terms on and off, since M_1 and M_2 are binary. The equation for FB expresses the idea that M_1 and M_2 act independently, so their chances of failure multiply. In the diagram, $\sigma = \tau = .1$, so $\Pr(FB = 1) = .9$ when either match acts alone, and .99 when they both act. Therefore, $\Delta_1 > 0$ for both M_1 and M_2. Both are strongly active (and promoters).

If, however, we redefine $FB = \max(s, t)$, then in the model shown, $\Pr(FB) = .9$ so long as any match is lit. Therefore, neither is strongly active. The original (deterministic) Matches is just a special case of this condition. So we need an account which considers the redundancy ranges.

2.2 HP2: Weakly active paths

Once again, the guiding intuition is that c was a token cause of e if there was a context where wiggling C made a difference. But now, we are allowed to consider more than just the actual context. We are allowed to consider all contexts in the *redundancy range*.

As before, let Δ be the difference in effect between c and c', but explicitly note the context $\mathbf{b}$. To determine whether a context is in the redundancy range, we must establish that changing context alone does not change the probability of e. Let Δ_2 be the difference between doing $C = c$ in the actual case ($\mathbf{b}$) and in the alternate case ($\mathbf{b}'$). That is:

$$\begin{aligned} \Delta_1 &= \Pr(e|I_{c,\mathbf{b}}) - \Pr(e|I_{c',\mathbf{b}}) \\ \Delta_2(x) &= \Pr(x|I_{c,\mathbf{b}}) - \Pr(x|I_{c,\mathbf{b}'}) \end{aligned}$$

We make the plausible assumption that the threshold for Δ_2 is the same ϵ as for Δ_1. (We will see later this has some nice consequences.) Then we can rewrite the definitions for Redundancy Range and Weakly Active Paths. Using $|x|$ for $\text{abs}(x)$:

Definition 10 (Redundancy range:) *Given an instantiated causal model $\mathcal{M}$ and a path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$, and the background $\beta = \Pi \backslash \phi = \{B_1, \ldots, B_m\}$, the set of values $\mathbf{b}' = \{b_1, \ldots, b_m\}$ lies in β's* redundancy range *for ϕ iff $|\Delta_2(x)| \leq \epsilon$ for all $x \in \phi$.*

Definition 11 (Weakly active:) *a path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$ is weakly active in causal model $\mathcal{M}$ iff $|\Delta_1| > \epsilon$ in any β-redundant model, where a β-redundant model replaces all $B \in \beta$ with values in their redundancy range for instantiated ϕ.*

So **HP2** says that if we can find such a weakly active path, c was token causally relevant. That is:

Definition 12 (HP2:) *Relative to c', c was token causally relevant to e iff for some ϕ and some $\mathbf{b}'$, $|\Delta_1| > \epsilon$ and $\forall x \in \phi : |\Delta_2(x)| \leq \epsilon$.*

Definition 13 (Token Promotion) *:* as before

Definition 14 (Token Prevention) *:* as before

2.3 HP3: Resetting Ranges

To extend **H3**, we need to generalize the idea of intervention. We shall need to be able to fix not only the *values* of a variable, but more generally, *probability distribution* over the values. For example, we intervene to freeze $SH = \{0.7, 0.3\}$. The details are described in Korb et al. (2004). We denote this kind of intervention as $I_{\vec{c}}$, denoting that we intervene on the whole distribution for C.

Usually we will intervene to fix the distribution of an unset variable at its *current values*. Such an intervention has no effect downstream, but cuts all back-paths, giving us exactly what we need: the ability to distinguish causes and effects by the asymmetries of intervention.

Recall that the main difference between **H3** and **H2** was that **H3** cares only about changes to E, not all of ϕ. So $\mathbf{b}'$ lies in β's *redundancy range* for ϕ iff $|\Delta_2(e)| \leq \epsilon$.

Then we can add the notion of resetting variables along ϕ. Let the **resetting ranges** $\mathcal{R}$ be all ways of intervening to set variables along ϕ to their *actual*

probability distributions. Then, $\forall \mathbf{r} \in \mathcal{R} : I_{\vec{\mathbf{r}}}$ we verify $|\Delta_2| \leq \epsilon$ remains true. (We either add $I_{\vec{\mathbf{r}}}$ to the conditionals in Δ_2, or consider the interventions in succession.) Therefore:

Definition 15 (HP3-active) *A path ϕ is active iff:*

1. *For some $\mathbf{b}'$ in the redundancy range for ϕ:*
   (a) $|\Delta_1| > \epsilon$. *(Wiggling C makes a difference.)*
   (b) $\forall \mathbf{r} \in \mathcal{R} : I_{\vec{\mathbf{r}}}$ *leaves* $|\Delta_2| \leq \epsilon$.

Then c was a token cause relative to c' iff there was such an active path from C to E.

2.4 Calculating Procedures

Since Bayesian networks are naturally computational objects, it is helpful to write down algorithms for calculating the relevant quantities. In fact, we would argue that it is necessary to prevent confusion over what is being held fixed, or tacitly assumed. Furthermore, if we are this explicit, it is straightforward to translate the theory into a computer program, and to compare various theories on a library of known cases. Finally, problems in the account can be treated as bugs in the algorithm, and modifications naturally seen as extensions to the existing account. Furthermore, future accounts can be easily compared.

We now present algorithms to determine whether paths are strongly or weakly active, and whether C was token causally relevant to E in accordance with **HP1** or **HP3**. Most algorithms presume the following are common knowledge:

C	(the Cause node)
E	(the Effect node)
net	(The network)
values	(The actual states of the variables up to E, inclusive.)
Π	(All ancestors of E.)
c'	(Optional contrast.)
ϵ	(Threshold. Defaults to 0.)

Furthermore, they presume the network is *uninstantiated* and that we have already performed basic error-checking, such as:

Error-checking: Raise an error in any of these cases:
(1) $\epsilon < 0$, or $\epsilon \geq 1$.
(2) c' is present but invalid (not a state in C).

In modern programming languages, this common knowledge could be implemented as a *class*. The main benefit is to reduce the number of variables we have to pass to our functions, and avoid duplication of error-checking. Some values (such as Π) can be calculated from the others.

The main functions are *HP1* and *HP3* (Figure 1) which test all paths between C and E to see if they are active. If any path is appropriately active, then C was token causally-relevant for E. Our implementations return the *most* active path, and its Δ_1. They both use *allpaths* (Figure 2) which finds all paths from

```
function HP1:
Presumes: as per strongly_active.

Find the most active path:
    let Δ = 0, φ_max = ∅
    foreach φ in allpaths(C, E):
        let d = strongly_active(φ, c')
        if |d| > |Δ|:
            let Δ = d, φ_max = φ
    return (Δ, φ)
```

```
function HP3:
Presumes: as per HP3_active.

Find the most active path:
    let Δ = 0, φ_max = ∅
    foreach φ in allpaths(C, E):
        let d = HP3_active(φ)
        if |d| > |Δ|:
            let Δ = d, φ_max = φ
    return (Δ, φ)
```

Figure 1: HP1 and HP3 return the most active of the weakly-active paths. If there are none, they return the tuple $(0, \emptyset)$.

```
function allpaths:
Variables:
    S      (source node)
    D      (destination node)

Φ = ∅ (List of all paths)
let kids = children of S
if D ∈ kids:
    add (S, D) to Φ
    remove D from kids
foreach kid ∈ kids:
    paths = allpaths(kid, D)
    foreach φ ∈ paths:
        append (S + φ) to Φ

return Φ
```

```
function strongly_active:
Uses delta() to see whether |Δ_1| > ε
Presumes: as per delta()
Variables:
    φ      The path ⟨C, Y_1, ..., Y_n, E⟩
    c'     Contrast class

let β = Π\φ    Background (B_1, ..., B_m)
return delta(C, β, c')
```

Figure 2: *allpaths* returns all (directed) paths from source (S) to destination (D), and *strongly_active* returns Δ_1 if the path is strongly active ($|\Delta_1| > \epsilon$), otherwise 0.

C to E by recursively searching down from C. Of course, they have different criteria for active, calling respectively, *strongly_active* (Figure 2) and *hp3_active* (Figure 3).

The function *strongly_active* is just a wrapper to *delta* (Figure 3), which does all the work. That deserves some discussion, because a symmetry lets us use a single function *delta* to compute either Δ_1 or Δ_2. First recall that in our case (omitting **HP2**), the quantities are:

$$\begin{aligned}\Delta_1 &= \Pr(e|I_{c,\mathbf{b}}) - \Pr(e|I_{c',\mathbf{b}}) \\ \Delta_2 &= \Pr(e|I_{c,\mathbf{b}}) - \Pr(e|I_{c,\mathbf{b}'})\end{aligned}$$

Notice we can write these as functions of two sets of variables x and y, notionally the cause c and background b:

$$\begin{aligned}\Delta_1(x,y) &= \Pr(e|I_{x,y}) - \Pr(e|I_{x',y}) \\ \Delta_2(x,y) &= \Pr(e|I_{x,y}) - \Pr(x|I_{x,y'})\end{aligned}$$

```
function HP3_active:
Presumes: as per strongly_active
save φ.
Variables:
    φ        ⟨C, Y_1, ..., Y_n, E⟩
    c'       Contrast class
let β = Π\φ      The Background

foreach b' in redundancy_ranges(β, φ):
    I_b'     Sets all vars in this b'.
    Check |Δ| > ε
    if not strongly_active (φ, c'):
        go to next b'

    Check ∀r ∈ R : I_r leaves |Δ_2| ≤ ε
    foreach r in resetting_ranges(φ):
        I_r   Freeze at actual distributions.
        if not redundant(b', φ):
            go to next b'
    Success! Found a b' that works!
    return strongly_active(φ, c')

    If we get here, nothing passed.
return 0
```

```
function delta:
Δ(X, Y, x') = Pr(e|I_{x,y}) − Pr(e|I_{x',y})
Presumes net is completely uninstantiated
Variables:
    X       The notional cause(s). (C or β)
    Y       The notional background. (β or C)
    x'      Contrast state(s) for X.
Error if length(X) ≠length(x')

Probability for actual state x, y:
    x = values(X),  y = values(Y)
    I_{x,y}
    update net
    let p_1 = belief of state e ∈ E

Probability for contrast class x', y:
    if x' is defined: I_{x'}
    else: unset X
    update net
    let p_2 = belief of state e ∈ E

unset(X, Y)     Return net to clean state.
let Δ = p_1 − p_2
if |Δ| > ε: return Δ
else: return 0
```

Figure 3: *HP3_active* returns Δ_1 if ϕ was active, and 0 otherwise. *delta* does the real work. Depending on how it is called, it calculates $|\Delta_1| > \epsilon$ or $|\Delta_2| \leq \epsilon$.

```
function redundancy_ranges:
Presumes net is completely uninstantiated
Variables:
    φ        (The path)

Get Pr(e|I_{c,b}).
    let β = Π\φ
    let b = values(β)
    I_{c,b}
    update net
    p_1 = Pr(e)

let RR = ∅
Try each possible instantiation of β.
    let B = all_instantiations(β)
    foreach full branch b' ∈ B:
        if redundant(b', φ):
            add b' to RR

unset c & b'     Clean network
return RR
```

```
function all_instantiations:
Variables:
    β        List of vars. The 'background'.

let B = pop(β)    Removes β_1
if length(β) = 0:    Base case
    return (states of B)    (a list)

Build A, a tree of states
    let A = ∅
    foreach state s ∈ B:
        add (s + all_instantiations(β)) to A

return A
```

Figure 4: *redundancy_ranges* returns a list of all redundant $\mathbf{b}'$. It uses *all_instantiations* to get $\mathcal{B}$, a tree showing all possible $\mathbf{b}'$.

function *redundant*:
Computes whether $|\Delta_2| \leq \epsilon$
Presumes: as per delta
Variables:
  $\mathbf{b}'$ *Contrast class for* β.
  ϕ *The path:* $\langle C, Y_1, \ldots, Y_n, E\rangle$
  Note: ensure length(b') = *length*(β).
  Generally some $b'_i = b_i$.

let $\beta = \Pi \backslash \phi$ *The background*
return *delta*(β, C, b')

function *resetting_ranges*:
Variables:
  ϕ *Path variables*

if ϕ is empty:
  return $\emptyset$ *Empty list*
let $p = \text{pop}(\phi)$ *Removes* ϕ_1
let s = current distribution of p
let rr = *resetting_ranges*(ϕ)
Make 2 branches: p set and p unset
  return [$[(p, s), rr]$, $[(p, -), rr]$]

Figure 5: *redundant* determines whether a set of background values b' is redundant for ϕ, and *resetting_ranges* finds all possible ways to freeze some variables along the active path at their actual distributions.

We can immediately see that:

$$\Delta_2(x, y) = \Delta(y, x)$$

Therefore to check that $|\Delta_2(c, \mathbf{b})| \leq \epsilon$, we merely swap c and $\mathbf{b}$, and check that $|\Delta(\mathbf{b}, c)| \not> \epsilon$. So both *strongly_active* and *hp3_active* (Figure 3) call *delta*.

HP3 (Figure 1) calls *hp3_active* on all paths, and once again returns the most active. However, our implementation of *hp3_active* is computationally intensive, so we might prefer that *HP3* return as soon as it finds *one* active path.

The function *hp3_active* in turn calls the utility functions *redundancy_ranges* and *resetting_ranges* (Figures 4 & 5). We have written these in the most straightforward way, with exhaustive search. In the toy problems common in the philosophy literature, this is fine, but for realistic networks, we should look for more efficient procedures that take advantage of blocking relations in the network.

Similarly, it has been easy to write these algorithms assuming it is easy and fast to intervene upon variables, or unset them. In practice, these steps may take time, and a programming implementation would likely modify the algorithms to reduce the number of such calls. For example, it might be faster to start with the network fully *instantiated*, and unset paths as required.

3 Examples

In which we show that **HP3** handles deterministic cases and also some novel probabilistic variants. In many cases the probabilistic variants are easier to handle, because paths are more likely to be strongly active.

Disjunctive Matches: As before, $FB = M_1 \vee M_2$. But this is equivalent to saying $\Pr(FB = 1) = M_1 \vee M_2$. Did $M_1 = 1$ cause $FB = 1$? Yes. $M_2 = 0$ is in the redundancy range, so we may set it. And then M_1 makes a difference to $\Pr(FB)$. More formally:

— Actually, $M_1 = 1$ and $FB = 1$.

— Let $\epsilon = 0$.

— Let $\mathbf{b}'$ denote $M_2 = 0$. Then $\Delta_2 = 0 \leq \epsilon$. So $\mathbf{b}'$ is in the redundancy range.

— $|\Delta_1| = 1 > \epsilon$. (There are no path variables to freeze.) So M_1 makes a difference.

By symmetry, the same holds for M_2.

Probabilistic Matches: As before, presume both matches cause FB independently, with failure probabilities α and β.

$$\begin{aligned} M_1 &= 1 \\ M_2 &= 1 \\ \Pr(FB) &= 1 - \sigma^{M_1}\tau^{M_2} \end{aligned}$$

Let $\sigma = .3$ and $\tau = .2$, so the probability of the forest burning down is:

M_1	M_2	$\Pr(FB = 1)$
0	0	0
0	1	.8
1	0	.7
1	1	.94

Then, letting $C = M_1$ and $\beta = M_2$, actually $c = b = 1$, and $\Pr(FB = 1) = .94$. Choosing $\epsilon = .1$:

1. Actually, c and e.
2. In the actual background ($M_2 = 1$):
   — The background change is redundant. (Trivially!)
   — There are no variables along ϕ, so no freezing.
   — If we set $M_1 = 1, M_2 = 0$, then $|\Delta_1| = .94 - .7 = .24 > \epsilon$.

The path is strongly active, and $M_1 = 1$ is an actual promoter (positively causally relevant). What about M_2? Proceeding similarly, we find that $|\Delta| = .94 - .8 = .14$, so $M_2 = 1$ is also an actual promoter. However, note that the answer is sensitive to ϵ. If we choose $\epsilon = 0.2$, then $M_1 = 1$ is a promoter, but $M_2 = 1$ is not relevant.

Also note that because Δ_1 and Δ_2 share the same ϵ, $M_x = 1$ is a cause iff it is not in the RR for M_y.

Max matches: The same structure as above, but $\Pr(FB = 1) = \max(sM_1, tM_2)$, where s is the probability of M_1 to cause $FB = 1$ when acting alone. Similarly for t.[6] (So $s = 1 - \sigma$.) Let $s = 0.8$ and $t = 0.7$.

First, consider $\epsilon < 0.1$. Then $M_1 = 1$ is a promoter of $FB = 1$, but $M_2 = 1$ is not. We have:

1. Actually, $M_1 = 1, M_2 = 1, FB = 1$, and $\Pr(FB = 1) = 0.8$.
2. The path from M_1 is strongly active: wiggling M_1 makes a difference, because $\Pr(FB = 1)$ changes from 0.7 to 0.8. $M_1 = 1$ is a promoter.
3. Wiggling M_2 makes no difference: $M_1 = 1$ is carrying the load.

[6] Here again we include M_x as a shortcut to an if... then statement. The shortcut only works when M_x are restricted to $\{0, 1\}$.

4. And $M_1 = 0$ is outside the RR for M_2, precisely because it qualifies as a cause all by itself. So M_2 is not causally relevant.

Now suppose $0.1 < \epsilon < 0.7$. Once again, $M_1 = 1$ is a promoter. But now, so is $M_2 = 1$.

1. Wiggling M_1 makes no noticeable difference. *However*, $M_2 = 0$ is now in the redundancy range for M_1.
2. Setting $M_2 = 0$, wiggling M_1 now makes $\Delta_1 = 0.8$. $M_1 = 1$ is a promoter.
3. Once again, wiggling M_2 makes no difference. However, $M_1 = 0$ is now in the redundancy range for M_2, because it only changes the probability by 0.1
4. Setting $M_1 = 0$, we now have that wiggling M_2 makes a difference: $\Delta_1 = 0.7$. $M_2 = 1$ is a promoter.

A note on ϵ: Obviously, we could set $\epsilon = 0.8$ and get that M_1 is a cause (promoter) and M_2 is not relevant. What are reasonable values for ϵ? We originally intended ϵ as a pragmatic concession: in practice we cannot insist $\epsilon = 0$ because we have finite data and limited computing precision. But as the previous example shows, it can also be a personal preference parameter. Indeed, determinists can always set $\epsilon \approx 1$. They will be blessed with few causes.

From here on, we shall assume $\epsilon = 0$ unless stated otherwise.

Bottle: As we saw above, in the deterministic models, we get the same answers as before: we just convert deterministic functions like $BT = SH \vee BH$ to probabilistic functions like $\Pr(BT = 1) = SH \vee BH$. Let's move on to probabilistic variants.

Probabilistic Bottle 1: Being more realistic, we allow that Suzy sometimes misses, and so does Billy. The model is $ST \xrightarrow{+} BS \xleftarrow{+} BT$. Here, both ST and BT are potentially causes, exactly as in Probabilistic Matches. If both Suzy and Billy have the same chance of hitting, then both count equally as promoters.

We do not want to accept that conclusion, but what exactly is wrong with it? We want to say, "But only Suzy's rock actually hit!" Suzy's throw is not a radium atom waiting to decay. There are other states which are relevant, and we know them to be easily observable. Our model omits the obviously important variable *Hit*: as Halpern & Pearl argued, we have the wrong model.

First, we know that Suzy throws faster, and will hit first. Second, we know that if her throw will hit, Billy's cannot. (This modelling assumption assumes she throws enough faster that the bottle will be shattered by the time Billy's rock arrives.) Third, we suspect that at some point prior to the bottle shattering, the outcome is sealed. That is, at some point, there is no more indeterminism in this system. Halpern & Pearl take this to be when one rock hits the bottle, presumably a very short time indeed before the bottle shatters.

Probabilistic Bottle 2: So we use Halpern & Pearl's model, but include the chance of failure. We add noise to the two initial arcs: $ST \xrightarrow{.7} SH$ and $BT \xrightarrow{.8} BH$. We leave the other arcs as they were, so: $SH \longrightarrow BS \longleftarrow BH$ and also $SH \longrightarrow BH$. Given knowledge only of ST and BT, the probability of the bottle smashing is:

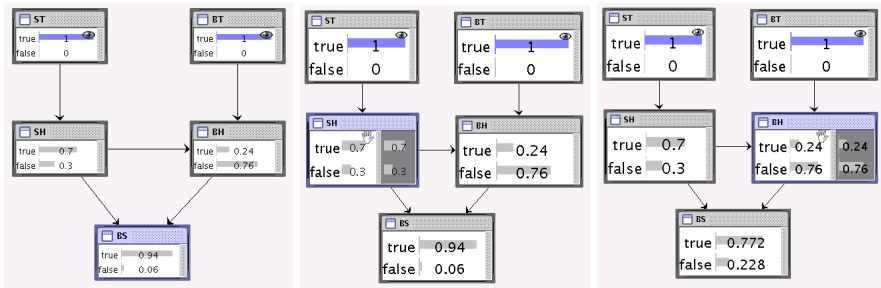


Figure 6: Probabilistic version of Bottle Smash: results of intervening to fix the probability distributions. Original state on the left. Freezing SH (middle) has no effect on BS. But freezing BH (right) does, so BH is not actually causally relevant.

ST	BT	$\Pr(BS = 1)$
0	0	0
0	1	.8
1	0	.7
1	1	.94

This is the same as for Matches. The redundancy ranges for both Suzy and Billy are degenerate. At first it seems that both $ST = 1$ and $BT = 1$ are promoters. However, we must allow for freezing the variables along the active path.

The network is shown in three states in Figure 6. When we freeze $SH = \{0.7, 0.3\}$, the distribution $\Pr(BS)$ is unchanged. $ST = 1$ is an actual cause. However, when we freeze $BH = \{.77, .23\}$, $\Pr(BS)$ changes, so $\Delta_2 > \epsilon$, meaning that BT is not actually a cause.

This works because interventions block the usual backwards flow of probability, priviledging upstream variables. (As an aside, we must be willing to consider the possibility that by intervention we can make both $SH = 1$ and $BH = 1$. Such constraints help to choose the proper models. We cannot, for example, replace BS with two binary variables $Smashed$ and $Intact$. (see Korb et al., 2005))

Probabilistic Boulder: In Hitchcock's example Boulder, suppose $F \xrightarrow{.9} D$, with the rest as before, including that the hiker won't duck spontaneously. Then everything proceeds as before, resulting in the same $\{0, 1\}$ probabilities at S that we had before.

Let's add some more noise. Suppose that the CPT at S is:

F	D	$\Pr(S = 1)$
0	0	1
0	1	1
1	0	.01
1	1	.8

Now, $F = 1$ is actually causally relevant. In fact, it hinders survival, as we might expect. Given that the hiker ducked (an intervention), $F = 1$ reduces the chances of survival relative to $F = 0$, as we can see in Figure 7. But that is just what we would expect.

Two Assassins: Hiddleston (forthcoming 2004) uses the example of the assassin and trainee to suggest that active-path accounts cannot handle probabilistic

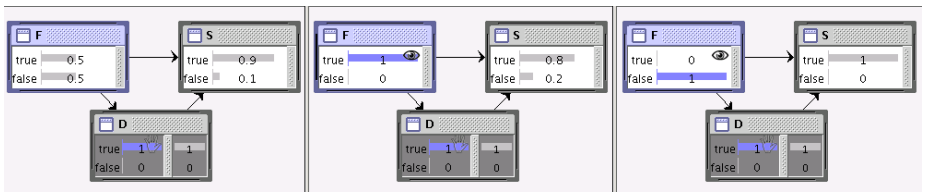
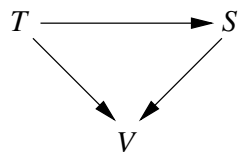


Figure 7: Probabilistic version of Boulder: now $F = 1$ is actually causally relevant. It decreases the chances of survival relative to $F = 0$ (and therefore relative to baseline).

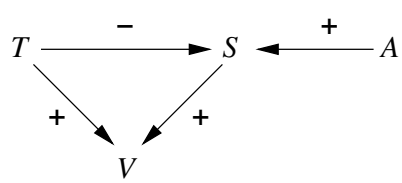
cases. The original story is that a trainee assassin T shoots at a victim V. The supervisor also takes aim, prepared to shoot if the trainee loses nerve. The supervisor is a crack shot.



$$\begin{aligned} T &= 1 \\ S &= \neg T \\ V &= \neg T \wedge \neg S \end{aligned}$$

In the actual case, the trainee shoots, the supervisor does not, and the victim dies. Although such examples are problems for older accounts of probabilistic relevance, it is easy for any account which considers one path at a time. The path $T \longrightarrow V$ is strongly active, so all accounts presented here get it right.

Hiddleston says that neither **H1** nor **H2** (and by implication **H3**) survives a probabilistic extension. Add an extra variable for whether the supervisor takes aim, and let the model be probabilistic, with no extreme probabilities. The supervisor is unlikely to fire without aiming, and always more likely to fire if he aims. Our model might look like this:



Consider whether $A = 1$ is an actual cause. The only background variable is T, and only the actual value $T = 1$ is redundant. In this context, the supervisor aiming ($A = 1$) changes the probability the victim will die, compared with not aiming ($A = 0$).[7] Freezing S at the distribution induced by $\{A, T\}$ has no effect on V, because $\{A, T\}$ already blocked the backwards flow of probability from S.

Therefore $A = 1$ is on an active path and is a cause. Hiddleston thinks this is wrong, because in actual fact the supervisor did not shoot.

[7] It happens to raise it. But what makes the path active is the fact that the probability changed at all, not that it went up. Hiddleston says we should "treat counterfactual dependence as probability raising" but that confuses relevance and role. His main criticism is independent of that point.

Now, on the standard "determination of probabilities" interpretation, changing the probability of an effect *is* causing that effect (either promoting or hindering). In fact, *that's all there is* to probabilistic causation. (Humphreys, 1989) Hence, we could argue that the original account yields the correct verdict in its domain.

But this is stretching things. Since S is in our model, it seems we ought to be able to account for the fact that $S = 0$. The problem is that our probabilistic extension must give credence to events ($S = 1$) that we know did not happen, making our path appear active when common sense says it was not.

So while we reject Hiddleston's claim that active path accounts cannot be made probabilistic, we agree with Hiddleston that these active-path accounts are limited by failing to attend to causal processes. Consider his deterministic Antidote example.

Antidote: Fred injects himself with harmless antidote ($A = 1$) on the false belief that he has been (or is about to be) poisoned (P). The model is simple: $P \longrightarrow S \longleftarrow A$, and $S = \neg P \vee A$. Fred survives when $S = 1$.

In fact there is no poison ($P = 0$), so the antidote was irrelevant. However, because $A = 1$, the RR includes $P = 1$. Had there been poison, antidote would have made a difference, hence the path $A \longrightarrow S$ was weakly active. As Hiddleston says, "That is the wrong answer. **H2** [and **H3**] counts ($A = 1$) an 'actual cause' merely because it *would have been* a cause if the circumstances had been different. . . ."

Hiddleston notes that **H1** correctly says ($A = 1$) is not a cause. But it also says ($P = 0$) is not a cause, leaving ($V = 1$) uncaused! That's the same problem we had with Matches. As with matches, if we make the example probabilistic, the difficulties vanish, and any of **HP1**, **HP2**, and **HP3** can solve it.

But we want an account that can handle this case as it stands. Hiddleston presents a solution in terms of Cheng-style causation. We offer our own solution using wounded arcs elsewhere (Korb et al., 2005).

$\boxed{\text{Does Hiddleston's solution solve } P = 0?}$

In passing we note that if we make this example probabilistic (with no extreme probabilities), then **HP3**

4 Conclusion

The active-path approach treats token causation as type causation in a restricted context. The trick is to find the proper context. We have extended the active-path analyses of deterministic token causation (Halpern and Pearl, 2001, 2004; Hitchcock, 2001) to probabilistic cases, defending that project against impossibility claims (Hiddleston, forthcoming 2004). Our solution makes use of a more general notion of intervention, allowing one to set an arbitrary probability distribution on a variable (Korb et al., 2004).

Nevertheless, we acknowledge some of Hiddleston's counterexamples, and agree with him that a full account must make reference to causal processes. We hope to present that account in a forthcoming piece.

5 Acknowledgements

A Monash Arts/IT grant for 2003/2004 supported this work. We also thank Chris Hitchcock, Lucas Hope, and students in our 2004 honours seminar on causation for useful discussion and criticism.

References

Halpern, J. Y., Pearl, J., 2001. Causes and explanations: A structural-model approach - part I: Causes. In: Proceedings of the Seventeenth Conference on Uncertainty in AI. pp. 194–202, superseded by BJPS 2004 Halpern and Pearl (2004).
URL `http://www.cs.cornell.edu/home/halpern/papers/actcaus.p%df`

Halpern, J. Y., Pearl, J., 2004. Causes and explanations: A structural-model approach - part I: Causes. British Journal for the Philosophy of Science ? (?), ?
URL `http://www.cs.cornell.edu/home/halpern/abstract.html#ac%tcaus`

Hiddleston, E., forthcoming 2004. Causal powers. British Journal for the Philosophy of Science ? (?), ?

Hitchcock, C. R., 2001. The intransitivity of causation revealed in equations and graphs. Journal of Philosophy XCVIII (6), 273–299.

Humphreys, P., 1989. The Chances of Explanation. Princeton University Press.

Korb, K., Hope, L., Nicholson, A., Axnick, K., 2004. Varieties of causal intervention. In: PRICAI'04.

Korb, K. B., Handfield, T., Oppy, G., Twardy, C., ? 2005. Causal reasoning with causal models. submitted to ? ? (?), ?

Lewis, D., 1973. Causation. The Journal of Philosophy 70, 556–67, reprinted (plus postscripts) in **?**, 159–240.