Token causation by probabilistic active paths

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

January 12, 2005

Abstract

We present a probabilistic extension to active path analyses of token causation (Halpern and Pearl, 2001, forthcoming; Hitchcock, 2001). 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 intervention variables, 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 (2005), 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 algorithmic formulation.

Causal models represent type causation, such as the general effect of smoking on lung cancer. Recently, Halpern & Pearl (2001; forthcoming) 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

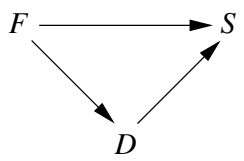
*Computer Science & Software Engineering, email:{ctwardy,korb}@csse.monash.edu.au
†Philosophy & Bioethics, email:{Graham.Oppy,Toby.Handfield}@arts.monash.edu.au

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, 2005) 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, we offer this account so that active path approaches may at least contend in the probabilistic domain. Significantly, we present algorithms for the theory (see Appendix), opening the way for automatically comparing accounts of causation on large sets of examples, rather than the usual handful of successes or failures presented in any one 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:



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

In this example, each variable is binary: 0 (false) or 1 (true). The first equation gives an initial condition: in the actual world, the boulder falls. The second says that the hiker ducks if 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. The causal model is fully specified by the graph, the equations, and the possible values for each variable.

The model encompasses all possibilities we are willing to entertain seriously, defining the set of acceptable counterfactuals, and providing truth conditions for evaluating them, as we see below.

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": had the hiker not ducked, the boulder must not have fallen, so ducking made no difference to survival. David Lewis ruled out backtracking in his original counterfactual analysis of causation (1973), because it would create an unholy symmetry between cause and effect.

Almost any non-backtracking account says that ducking saved the hiker. Causal models are non-backtracking because we represent counterfactuals by *interventions* rather than observations. Instead of treating $D = 1$ as an observation on the system, 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 accords 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 other 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 (APTC) *$C = c$ was a token cause of $E = e$ iff there was an **active path** from C to E (and actually $C = c$ and $E = e$).*

A **path** is a walk from C to E along the arrows, in the direction of the arrows. Intuitively, a path is active if it makes a difference. We take the simple case first.

A path ϕ 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. (In fact they are not active at all.) 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 $\langle F, D, S\rangle$. 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).

In general, paths can be long chains like $C \longrightarrow X \longrightarrow Y \longrightarrow E$. We want to leave exactly one path ϕ open, such as the path through D, and block all others.[1] 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 we use Π for all the ancestors. Then we define:

Definition 2 (Background β) *Given a target variable E, and a path ϕ ending with E, the background β is all the ancestors of E except for those in ϕ, which we write: $\beta = \Pi \backslash \phi$.*

Definition 3 (β-actual model) *Given causal model $\mathcal{M}$ with target variable E, the β-actual model replaces all the equations for β with their actual values. The replacements are interventions.*

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 $I_{C=c}$ or just 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 $C = c'$. (Similarly, in later sections we may write "set c'" to mean "set $C = c'$", using lowercase letters to denote states of the variable with the same uppercase letter.)

Then we define:

Definition 4 (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.*

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.

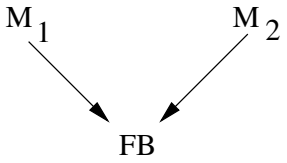
[1] For simplicity, assume any descendants of our target variable (E) are unknown. If they were known, they would inform us about E, obscuring the causal impact of C.

If we limit active paths to *strongly* active paths, APTC gives us Hitchcock's simple account of causation. Following Hiddleston (2005), we shall call this **H1**.

Unfortunately, this simple account cannot handle symmetric overdetermination nor late cutting.

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 4, 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 on their own make no difference to the result, or to other variables on our chosen path ϕ (here $\langle M_1, FB\rangle$). Hitchcock calls this set of values the **redundancy range** for ϕ. Let us set $M_2 = 0$, a value in the redundancy range. We immediately see the hidden dependency: now M_1 makes a difference! The intuition is that M_1 was really contributing to the causal history, but that its effect was accidentally masked.

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

Definition 5 (Redundancy range) *Given an instantiated causal model $\mathcal{M}$, a path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$, and the background $\beta = \Pi\backslash\phi = \{B_1, \ldots, B_m\}$, then the corresponding vector of values $\mathbf{b}' = \langle b_1, \ldots, b_m\rangle$ 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 background variables so that they make no difference to anything in ϕ. Note that the actual values of β are also in the redundancy range!

Definition 6 (β-redundant model) *Given a causal model $\mathcal{M}$ with background β, a β-redundant model replaces all $B \in \beta$ with values in their redundancy range for instantiated ϕ. The replacements are interventions.*

An active path is a path that becomes strongly active when we set background variables to values in their redundancy range. Since the redundancy range includes the actual values, strongly active paths are themselves active.[2]

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

Definition 7 (Active path) *A path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$ is active in causal model $\mathcal{M}$ iff I_c makes a difference (relative to $I_{c'}$) to E in some β-redundant model.*

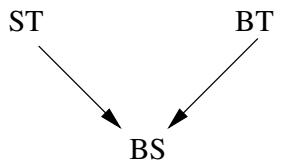
Using active paths Hitchcock's definition can now handle cases of symmetric overdetermination: both paths in Matches are active (though not strongly active), so both matches are causes according to Definition 1. Hiddleston calls this account **H2**.

1.3 Late Cutting

Late cutting is a kind of pre-emptive 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.



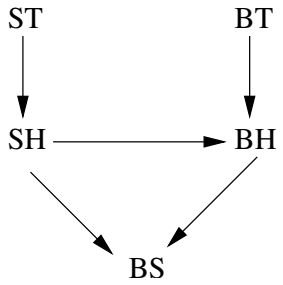
[2]For this reason we do not follow Hitchcock in calling it "weakly active".

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

This is the same model as Disjunctive Matches, so it gives the same answer: both Billy's throw and Suzy's throw are 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 the verdict is unchanged 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.

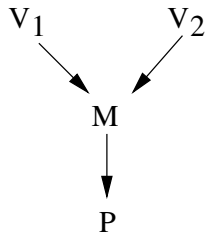
But Billy's path ($BT \longrightarrow BH \longrightarrow BS$) is not active. It is not active because it is not strongly active and the redundancy range is degenerate. It is not strongly active because Suzy's rock actually hits, so $BH = 0$ no matter what Billy does. The redundancy range is degenerate because only ($SH = 0$) leaves BH unchanged, and BH is on our path.

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 (forthcoming) present one which **H2** cannot handle.[3]

1.4 Problem and diagnosis

The Voting Machine example (Halpern and Pearl, forthcoming, Example A.3) shows that **H2** is too strict about leaving ϕ unchanged. 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:



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

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 active at all, 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.

This is Halpern & Pearl's account:[4]

[3] Our thanks to Chris Hitchcock for directing our attention to this case.

[4] We omit the exogenous error terms, which serve mostly to complicate the expressions and assuage deterministic leanings. Second, we use helpful variable names. Finally, we presume C is a single variable, so do not need their minimality condition AC3.

Definition 8 (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.*

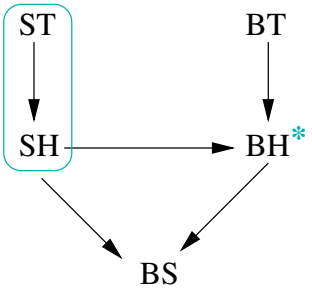
Recall that "set b" is shorthand for "set $B = b$".

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.

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 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.

In the hope that it will help readers, we present this solution as two opposing players taking turns in a short game. The "advocate" tries to show that some $C = c$ is a cause of $E = e$.[5] To this end, he nominates an active path, a value $c' \neq c$, and context b' such that given b', c yields

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

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).

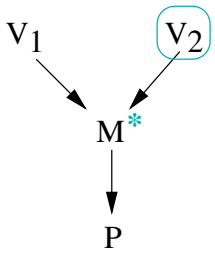
In the augmented bottle-smashing model, the dialogue would go something like this:

ADV: $(BT = 1)$ $[c]$ is a cause of $(BS = 1)$ $[e]$. The path is $\phi = BT \longrightarrow BH \longrightarrow BS$. Background: $\beta = \{ST, SH\}$, and b' sets $SH = 0$. Then the cause makes a difference: c gives e, but c' $[BT = 0]$ gives e' $[BS = 0]$.

NAY: I agree that $(SH = 0)$ is in the redundancy range for ϕ. But since BH is on the original path, I reset it to its original value, $BH = 0$. Now your cause makes no difference.

ADV: I concede.

In the voting machine case, it would go something like this:



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

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.

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

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 of β allows c to make a difference. In the deterministic case, if fixing some subset works, then there will also be a way to fix all of β. In the probabilistic case, that won't be true.

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 APTC: c was a token cause (now meaning token causally *relevant*) only if 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.

Following Hitchcock, our first version uses strongly 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 simply means that there is a change in probabilities rather than outcomes, so it is worth rewriting. Because we shall later introduce another kind of difference making, we shall call this Δ_1 rather than just Δ. Define

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

Intuitively, a path is strongly active when $\Delta_1 \neq 0$. However, with finite data and limited computing power, we cannot be so stringent. Let $0 \leq \epsilon < 1$ denote some tolerance threshold below which we consider the change in probability to be insignificant. More precisely, we set ϵ to a value which balances the disutility of false positives with those of false negatives. Then:

Definition 9 (Strongly active) *A path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$ is strongly active in causal model $\mathcal{M}$ iff $|\Delta_1| > \epsilon$ in the β-actual model.*

Recall that replacing the background β is an intervention, so the equation for Δ_1 implicitly conditions on $I_\beta = I_{b_1,\ldots,b_m}$. As before, this definition requires an explicit contrast class c'.[6]

[6]Because we have *intervened* to set β, we can let c' default to the distribution on C after fixing β.

We have chosen to make Δ_1 relative to the probability of the actual state e, because we think this is the most straightforward extension of the deterministic accounts. However, if E has more than two states, it is possible that c versus c' changes the distribution over E without affecting a particular state e. For example, suppose E lists several possible causes of death. For a given background, it may well be that pipe smoking gives the same probability of fatal heart attack as cigarette smoking, but that pipe smoking makes you less likely to die of lung cancer (and more likely you will die of esophogeal cancer or natural causes).

If we want such changes to count, we could just measure the distance between the entire distribution $\Pr(E|I_c) - \Pr(E|I_{c'})$, using say Kullback-Leibler "distance" or Bhattacharyaa distance. We leave that aside for this paper.

Roughly then, **HP1** says:

Definition 10 (HP1) *c was token causally relevant to e iff there was a strongly active path, and actually $C = c$ and $E = e$.*

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 C may both promote and prevent E. (Of course, in any particular context the *net* effect will be either positive, negative, or neutral.)

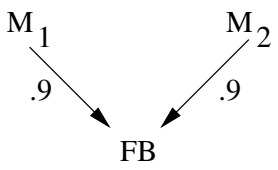
Definition 11 (Token Promotion) *c token promotes e iff for some path ϕ:*

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 12 (Token Inhibition) *c token inhibits 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 $(1 - s)$ and $(1 - t)$. Assume the forest cannot burn down spontaneously. Then if $s = t = .9$, and the matches act independently, our model might be:



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

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, $s = t = .9$, 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 — one with extreme probabilities.

2.2 HP2: 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_1 be the difference in effect between c and c', but relative to the context $\mathbf{b}$. To determine whether a context is in the redundancy range, we must establish that changing context alone leaves unchanged the probabilities of all actual states along our path, including e. Let x range over these actual states of variables along our path. Let Δ_2 be the difference between doing c in the actual case ($\mathbf{b}$) and in the alternative 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. Then we can rewrite the definitions for Redundancy Range and Active Paths:

Definition 13 (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 vector of values $\mathbf{b'} = \langle b_1, \ldots, b_m\rangle$ lies in β's* redundancy range *for ϕ iff for all variables $X \in \phi$ with actual state x, $|\Delta_2(x)| \leq \epsilon$.*

Definition 14 (β-redundant model) as before

Definition 15 (Active) *A path $\phi = \langle C, Y_1, \ldots, Y_n, E\rangle$ is active in causal model $\mathcal{M}$ iff $|\Delta_1| > \epsilon$ in some β-redundant model.*

So **HP2** says that if we can find such an active path, c was token causally relevant:

Definition 16 (HP2) *c was token causally relevant to e iff there was an active path and $C = c$ and $E = e$.*

This happens iff for some ϕ and some $\mathbf{b}'$, $|\Delta_1| > \epsilon$ and $\forall x \in \phi : |\Delta_2(x)| \leq \epsilon$.

Definition 17 (Token Promotion) *as before*

Definition 18 (Token Inhibition) *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 "clamp" not only the *values* of a variable, but more generally, *probability distribution* over the values. For example, we might clamp $SH = \langle 0.7, 0.3 \rangle$, rendering SH insensitive to other influences. The details are described in Korb et al. (2004). (We ignore their provision for imperfect or leaky clamps which still allow the original parents to influence the clamped variable.) We denote this kind of intervention as $I_{\vec{c}}$, meaning that we provide a new distribution for C.

Usually we will clamp 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 $|\Delta_2(e)| \leq \epsilon$. We can just write Δ_2 to mean $\Delta_2(e)$, since we no longer need to range over all $X \in \phi$.

Then we can add the notion of resetting variables along ϕ, by clamping them to their *actual* probability distributions. Any variable along ϕ can be clamped or not, so if there are n such variables, there are 2^n possible combinations. Let $\mathcal{R}$ be the set of all 2^n combinations.[7] Then, $\forall \mathbf{r} \in \mathcal{R}$ we intervene $I_{\vec{\mathbf{r}}}$ and verify $|\Delta_2| \leq \epsilon$ remains true. (We add our clamp $I_{\vec{\mathbf{r}}}$ to the interventions already done for Δ_2.) Therefore:

Definition 19 (HP3-active) *A path ϕ is active iff for some $\mathbf{b}'$ in the redundancy range for ϕ:*

1. *$|\Delta_1| > \epsilon$. (Wiggling C makes a difference.)*
2. *$\forall \mathbf{r} \in \mathcal{R}$, $I_{\vec{\mathbf{r}}}$ 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, and actually $C = c$ and $E = e$.

[7] In parallel with "redundancy ranges", we can call $\mathcal{R}$ the "resetting ranges".

2.4 An anti-symmetry proof

We show that when Δ_1 and Δ_2 have the same ϵ, then a cause is just the opposite of a redundant background condition. The intuition is clear: a cause is something which makes enough of a difference, and a redundant background condition is something which does not. The proof is straightforward, provided we can extend the notion of "cause" to multi-variable sets, even if only analogously.

We first write both Δ_1 and Δ_2 as functions of a common abstract Δ. All practical consequences flow from that lemma, regardless whether you wish to restrict the word "cause" to single variables. We assume all variables are separately intervenable, which entails they are not logically or mathematically related. We do the proof for **HP3**.

Lemma $\Delta_1(x, y, z) = \Delta_2(y, x, z)$
Proof. In **HP3**, Δ_2 does not range over all $X \in \phi$, so the quantities are simply:

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

Notice we can write both Δ_1 and Δ_2 as variants of a single function Δ:

$$\Delta(x, y, z) = \Pr(e|I_{x,y}) - \Pr(e|I_{z,y})$$

where notionally, x is the cause, y is the actual background, and z is the contrast class. Then:

$$\begin{aligned} \Delta_1 &= \Delta(c, \mathbf{b}, c') \\ \Delta_2 &= \Delta(\mathbf{b}, c, \mathbf{b}') \end{aligned}$$

or more generally,

$$\begin{aligned} \Delta_1(x, y, z) &= \Delta(x, y, z) \\ \Delta_2(x, y, z) &= \Delta(y, x, z) \end{aligned}$$

We see immediately that: $\Delta_1(x, y, z) = \Delta_2(y, x, z)$.

Theorem 1 (Anti-symmetry) *c (vs c') token causes e in context $\mathbf{b}'$ iff $\mathbf{b}'$ (vs the actual background $\mathbf{b}$) does* not *cause e in context c.*

Proof. Suppose that actually $c, e, \mathbf{b}$.

Suppose c (vs c') token causes e. Then there must be a redundant background $\mathbf{b}'$, meaning that $|\Delta_2| \leq \epsilon$. By substitution from the Lemma,

$$|\Delta(\mathbf{b}, c, \mathbf{b}')| \leq \epsilon \tag{1}$$

Now suppose for contradiction that same redundant context $\mathbf{b}'$ was itself a "cause" of e, in context c. (The context is trivially redundant, since actually c.) Then $|\Delta_1| > \epsilon$. By the Δ Lemma, that means:

$$|\Delta(\mathbf{b}, c, \mathbf{b}')| > \epsilon \tag{2}$$

which directly contradicts Equation 1.

The proof in the other direction merely reverses the inequalities. QED

Corollary: The claim also holds for **HP1**. This follows immediately because **HP1** just requires that the redundant background $\mathbf{b}'$ is the actual background $\mathbf{b}$. Therefore both 1) and 2 are 0 trivially, giving us the contradiction.

Corollary: The claim also holds for **HP2**. **HP2** requires that $|\Delta_2(x)| \leq \epsilon$ for all actual values x of variables along the active path. Therefore it holds for $|\Delta_2(e)|$, which is what we used in Equation 1.

3 Examples

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.

Deterministic 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, thus we may set it so. And then M_1 makes a difference to $\Pr(FB)$. More formally:

— Actually, $M_1 = 1$ and $FB = 1$.

— Let $\epsilon = 0$.

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

— $|\Delta_1| = 1 > \epsilon$. (There are no path variables to clamp.) 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 σ and τ.

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

For example, 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, since we have chosen the actual value for M_2!)
   - There are no variables along ϕ, so no clamping is needed.
   - $|\Delta_1| = |\Pr(e|I_{M_1=1,M_2=1}) - \Pr(e|I_{M_1=0,M_2=1})| = |.94 - .8| = .14 > \epsilon$.

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 $|\Delta_1| = |.94 - .7| = .24$, so $M_2 = 1$ is also an actual promoter. However, note that the answer is sensitive to ϵ. Depending on how strong we require a causal factor to be, we can exclude one or both of M_1 and M_2. For example, if we choose $\epsilon = 0.2$, then $M_2 = 1$ is a promoter, but $M_1 = 1$ is not relevant.

Also note that by Theorem 1, $M_x = 1$ is a cause iff it is not a suitable background for making M_y a cause.

Max matches: This parallels Probabilistic Matches, but we let $\Pr(FB = 1) = \max(sM_1, tM_2)$. As before, M_x is a binary indicator variable for the proposition "match x is lit", s is the probability of FB given $M_1\&\neg M_2$, and similarly for t. Let $s = 0.7$ and $t = 0.8$.

First, consider $\epsilon < 0.1$. Then $M_2 = 1$ is a promoter of $FB = 1$, but $M_1 = 1$ 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_2 is strongly active: wiggling M_2 makes a difference, because $\Pr(FB = 1)$ changes from 0.8 to 0.7. $M_2 = 1$ is a promoter.
3. Wiggling M_1 makes no difference: $M_2 = 1$ is carrying the load.
4. And $M_2 = 0$ is outside the redundancy range for M_1, precisely because it makes a difference all by itself. So M_1 is not causally relevant.

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

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: As the previous example shows, ϵ is not just a concession to limited precision; it could also be a personal preference parameter based upon the disutility of error. Indeed, deterministes can always set $\epsilon \approx 1$. They will be blessed with few causes.

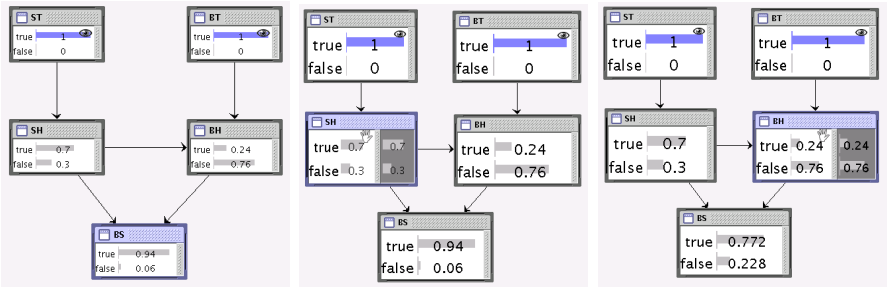
From here on, we shall assume $\epsilon = 0$ unless stated otherwise.

Deterministic 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 $ST \longrightarrow BS \longleftarrow BT$. 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.

Yet we want to say, "But only Suzy's rock actually hit!" As Halpern & Pearl argued, we have the wrong model: we have omitted the obviously important variable *Hit*. Given the story, the probability Billy hits given that Suzy throws is less than the probability that Suzy hits, given that Billy throws.

First, we know that Suzy throws faster, and will hit first, if at all. 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.



(a) (b) (c)

Figure 1: Probabilistic version of Bottle Smash: results of intervening to fix the probability distributions. (a) the original state; (b) clamping SH has no effect on BS; (c) but clamping BH does

Probabilistic Bottle 2: So we use Halpern & Pearl's model (Figure 1, but include the chance of failure. We add noise to the two initial arcs, for example: $ST \xrightarrow{.7} SH$ 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:

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 clamping the variables along the active path.

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

This works because these interventions block the usual backwards flow of probability, privileging 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: once the boulder has fallen (or not), everything is determined, so F is not actually causally relevant to S.

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 inhibits survival, as we might expect. Given that the hiker ducked, $F=1$ reduces the chances of survival relative to $F=0$, as we can see in Figure 2. But that is just what we would expect.

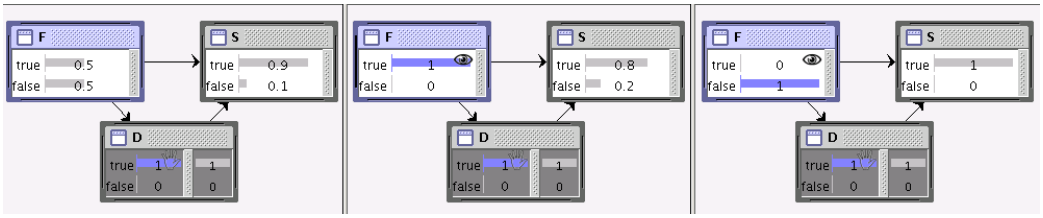
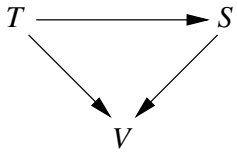


Figure 2: Probabilistic version of Boulder: now $F=1$ is actually causally relevant. It decreases the chances of survival relative to $F=0$ (and therefore relative to baseline).

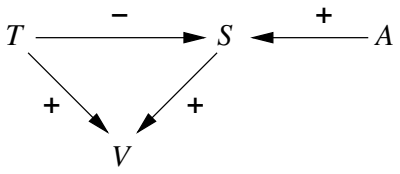
Trainee Assassin: Hiddleston (2005) uses the example of the assassin and trainee to suggest that active path accounts cannot handle probabilistic cases. The original story is that a trainee assassin T shoots at a victim V. The supervisor also takes aim, prepared to shoot if the trainee loses nerve. Being deterministic, Victim's survival requires neither shot to be fired.



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

In the actual case, the trainee shoots, the supervisor does not, and the victim dies. Although such examples are problems for simple dependency accounts, 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.

Stochastic Assassin: Hiddleston says that the analysis from either **H1** or **H2** (and by implication **H3**) does not survive a probabilistic extension. Add an extra variable for the supervisor taking 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 ($A = 0$).[8] Clamping S at the prior distribution induced by $\{A, T\}$ has no effect on V.

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.

But note that we have asked if $A = 1$ is a cause *in the null context*. We pretend we do not know that $S = 0$. And here, on the standard "determination of probabilities" interpretation, changing the probability of an effect *is* causing that effect (either promoting or inhibiting). In fact, *that's all there is* to probabilistic causation (Humphreys, 1989). Hence, we could argue that A *is* a cause.

Our "common sense" appears to be taking advantage of the causal foreground, namely, that $S = 0$. Fixing ($S = 0$) cuts the active path and renders A noncausal. In that context we get the right answer: $A = 1$ was not a cause because in fact, $S = 0$.

However, our account succumbs to Hiddleston's 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).[9] The model is simple: $P \dashrightarrow S \dashleftarrow A$, 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 redundancy range includes $P = 1$. Had there been poison, antidote would have made a difference, hence the path $A \dashrightarrow S$ was active. As Hiddleston says, "That is the wrong answer. **H2** [and **H3**] counts ($A = 1$) an

[8] 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.

[9] Or if you prefer, the king's bodyguard puts antidote in the king's coffee, but in fact there was no poison there to neutralize.

'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 $(S = 1)$ 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, even deterministically.

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). The account differs from Hiddleston's, but we agree that the accounts presented here are limited by failing to attend to causal processes.

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, forthcoming; Hitchcock, 2001) to probabilistic cases, defending that project against impossibility claims (Hiddleston, 2005). 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 begin to develop such an account in Korb et al. (2005).

Appendix: Calculating Procedures

Since Bayesian networks are 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 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)

function *HP1*:
Presumes: as per strongly_active.

Find the most active path:
let $\Delta = 0,\ \phi_{max} = \emptyset$
foreach ϕ in $allpaths(C, E)$:
let $d = strongly\_active(\phi, c')$
if $|d| > |\Delta|$:
let $\Delta = d,\ \phi_{max} = \phi$
return (Δ, ϕ)

function *HP3*:
Presumes: as per HP3_active.

Find the most active path:
let $\Delta = 0, \phi_{max} = \emptyset$
foreach ϕ in $allpaths(C, E)$:
let $d = HP3\_active(\phi)$
if $|d| > |\Delta|$:
let $\Delta = d, \phi_{max} = \phi$
return (Δ, ϕ)

Figure 3: HP1 and HP3 return the most active of the active paths. If there are none, they return the tuple $(0, \emptyset)$.

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 object-oriented 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 3) 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 (largest $|\Delta_1|$), and its Δ_1. They both use *allpaths* (Figure 4) which finds all paths from C to E by recursively searching down from C. Of course, they have different criteria for active, calling respectively, *strongly_active* (Figure 4) and *hp3_active* (Figure 5).

The functions *strongly_active* (Figure 4) and *hp3_active* (Figure 5) are just front-ends to *delta* (Figure 5), which does all the work, using the symmetry discussed in Section 2.4.

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

$\Phi = \emptyset$ *(List of all paths)*
let *kids* = children of S
if $D \in$ *kids*:
add (S, D) to Φ
remove D from *kids*
foreach $kid \in kids$:
$paths =$ allpaths(kid, D)
foreach $\phi \in paths$:
add $(S + \phi)$ to Φ

return Φ

function *strongly_active*:
Uses delta() to see whether $|\Delta_1| > \epsilon$
Presumes: as per delta()
Variables:
ϕ *The path* $\langle C, Y_1, \ldots, Y_n, E \rangle$
c' *Contrast class*

let $\beta = \Pi \backslash \phi$ *Background* $(B_1, \ldots, B_m)$
return *delta*(C, β, c')

Figure 4: *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.

HP3 (Figure 3) calls *hp3_active* on all paths, and once again returns the most active. However, *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 6 & 7). 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, these algorithms assume 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.

function *HP3_active*:
Presumes: as per strongly_active save ϕ.
Variables:
ϕ $\langle C, Y_1, \ldots, Y_n, E\rangle$
c' *Contrast class*
let $\beta = \Pi \backslash \phi$ *The Background*

foreach $\mathbf{b}'$ in *redundancy_ranges*(β, ϕ):
$I_{\mathbf{b}'}$ *Sets all vars in this* $\mathbf{b}'$.
Check $|\Delta| > \epsilon$
if not *strongly_active* (ϕ, c'):
go to next $\mathbf{b}'$

Check $\forall \mathbf{r} \in \mathcal{R} : I_{\vec{\mathbf{r}}}$ *leaves* $|\Delta_2| \leq \epsilon$
foreach r in *resetting_ranges*(ϕ):
$I_{\vec{\mathbf{r}}}$ *Clamp at actual distributions.*
if not *redundant*$(\mathbf{b}', \phi)$:
go to next $\mathbf{b}'$
Success! Found a ***b'*** *that works!*
return *strongly_active*(ϕ, c')

If we get here, nothing passed.
return 0

function *delta*:
$\Delta(X, Y, x') = \Pr(e|I_{x,y}) - \Pr(e|I_{x',y})$
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$(X) \neq$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 $e \in E$

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

unset(X, Y) *Return net to clean state.*
let $\Delta = p_1 - p_2$
if $|\Delta| > \epsilon$: **return** Δ
else: **return** 0

Figure 5: *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 6: *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 |Δ_2| ≤ ε
Presumes: as per delta
Variables:
    b'      Contrast class for β.
    φ       The path: ⟨C, Y_1, ..., Y_n, E⟩
    Note: ensure length(b') = length(β).
    Generally some b'_i = b_i.

let β = Π\φ    The background
return delta(β, C, b')
```

```
function resetting_ranges:
Variables:
    φ       Path variables

if φ is empty:
    return ∅    Empty list
let p = pop(φ)    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 7: *redundant* determines whether a set of background values b' is redundant for ϕ, and *resetting_ranges* finds all possible ways to clamp some variables along the active path at their actual distributions.

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 (forthcoming).
URL `http://www.cs.cornell.edu/home/halpern/papers/actcaus.p%df`

Halpern, J. Y., Pearl, J., forthcoming. 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., 2005. Causal powers. British Journal for the Philosophy of Science 56, 27–59.

Hitchcock, C. R., 2001. The intransitivity of causation revealed in equations and graphs. Journal of Philosophy 98 (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: Pacific Rim Conference on AI'04. pp. 322–31.

Korb, K. B., Handfield, T., Oppy, G., Twardy, C., 2005. Causal reasoning with causal models.

Lewis, D., 1973. Causation. The Journal of Philosophy 70, 556–67, reprinted (plus postscripts) in Lewis 1986, 159–240.

Lewis, D., 1986. Philosophical Papers. Vol. 2. Oxford University Press, New York.