Actual causation by probabilistic active paths
Online Supplement

Charles R. Twardy and Kevin B. Korb[†‡]

Abstract

This online supplement contains peer-reviewed material that would not fit in the PSA 2010 Proceedings article, Twardy and Korb (2011). That article presents a probabilistic extension of active-path analyses of actual causation. The extension uses "soft" interventions (Korb et al., 2004) as the analog of resetting path variables to their actual value. Soft interventions allow the "actual value" to be a probability distribution. The resulting account can handle at least as wide a range of examples as the original accounts, without assuming determinism.

A Additional Detail

Figure for Bottle example (Section 2)

Figure 1 accompanies the following text, which appeared in the Bottle example at the end of Section 2, just after Definition 8.

> In the extended bottle-smashing model, Suzy's throw $ST = 1$ remains an actual cause of $BS = 1$ using a redundancy range to reveal the dependency. Choosing $\{BT, BH\}$ as our background, and setting $BH = 0$, we see that ST makes a difference. Resetting has no effect because actually $SH = 1$.

[†]Charles R. Twardy, C[4]I Center, George Mason University, 4400 University Dr., MS 4B5, Fairfax, VA 22030; e-mail: ctwardy@c4i.gmu.edu; Kevin B. Korb, School of Info Tech, Monash University, Clayton, VIC 3800, Australia; e-mail: kbkorb@gmail.com.

[‡]This work was funded in part by a Monash Arts/IT grant. We are grateful for comments by Toby Handfield, Graham Oppy, Lucas Hope, Chris Hitchcock, an anonymous reviewer, and participants in the Monash 2004 causation seminar as well as PSA 2010 in Montréal.

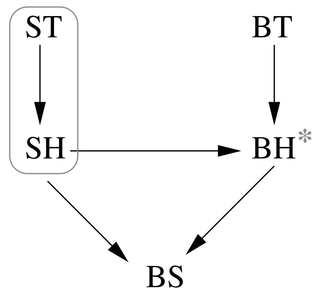
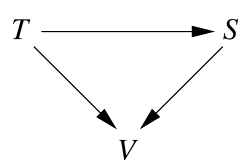


Figure 1: Modified bottle example showing the background in a box, and the rest variable with an asterisk. BH = Billy hits; BS = bottle shatters; BT = Billy throws; SH = Suzy hits; ST = Suzy throws.

Figure for The Trainee Assassin (Section 4.2)

Figure 4 shows the model for the deterministic version of Trainee Assassin presented in Section 4.2. Recall that Trainee T shoots at a victim V, but Supervisor will shoot if Trainee loses nerve. Being deterministic, Victim survives only when neither shot is fired. In the actual case, the trainee shoots, the supervisor does not, and the victim dies. As with Boulder, there is no marginal effect, but any path-based account works, because the path $T \longrightarrow V$ is strongly active.



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

Figure 2: Trainee Assassin example. T = Trainee shoots, S = Supervisor shoots, V = Victim survives.

B The Background is Not a Cause

The redundancy range analysis of Hitchcock amounts to asserting that the background is not a cause, which leads to a unified notation for treating cause and context.

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

Define Δ:

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

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

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

If c (vs c') is an actual cause of e, there must be a redundant background $\mathbf{b}'$ where c vs c' makes a difference. Asserting redundancy amounts to saying the background $\mathbf{b}'$ vs actual background $\mathbf{b}$ is not itself a cause:

$$|\Delta_2| = |\Delta(\mathbf{b}, \mathbf{b}', c)| \leq \varepsilon$$

The condition is trivially satisfied if $\mathbf{b} = \mathbf{b}'$. An implication is that if c is a cause, ε is constrained by the relative probabilities of c vs c' and $\mathbf{b}$ vs $\mathbf{b}'$.

C Probabilistic Boulder

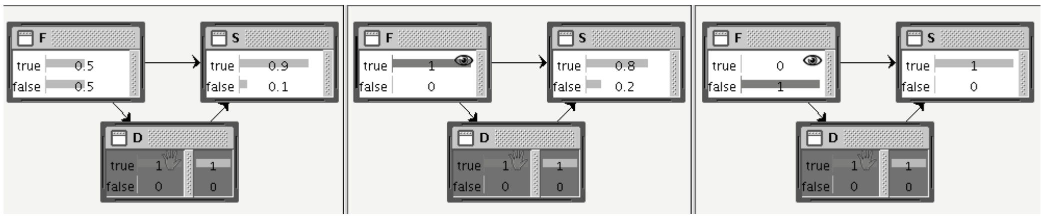


Figure 3: Probabilistic version of Boulder, where Hiker Survives (S=1) is stochastic. Here, Boulder Falls (F=1) is also an actual cause. Holding fixed that Hiker Ducks (D=1), F=1 decreases the chances of survival relative to F=0.

We can deal with the original Probabilistic Boulder without resetting. As before, D is an actual cause. If we suppose $F \xrightarrow{.9} D$, then everything proceeds as before, and D makes a difference; but F does not make a difference, because the direct arc $F \longrightarrow S$ is inactive (by setting $\mathbf{b}' : D = 1$), and we have no background variable that could block $F \longrightarrow S$ and isolate $F \longrightarrow D \longrightarrow S$. However, if we add noise at S, so that ducking no longer guarantees survival, we find an influence for F. Suppose that the conditional probability table (CPT) at S is:

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

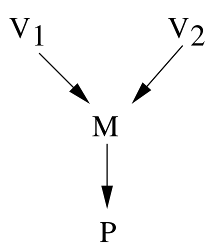
Given that the hiker ducked, $F = 1$ reduces the chances of survival relative to $F = 0$, as we can see in figure 3. But that is just what we would expect.
Of course, $F = 1$ is not an actual cause of $S = 1$. It is rather a counterfactual cause of $S = 0$ —i.e., a counterfactual *preventer* of $S = 1$. By always requiring that *actually* $E = e$ ($S = 1$ in this case), the accounts of Halpern & Pearl and Hitchcock deny F in this case the status of actual cause. PAP, however, finds it to be active, which we consider to be more appropriate. The awkwardness of calling a Falling Boulder an actual cause of Survival, when its *activity* all pulls in the opposite direction, is handled in English by such locutions as: "The hiker survived *despite* the boulder falling." The existence of such locutions supports,

rather than undermines, the status of falling boulders as actual causes in such cases. The inclusion of the clause requiring for actual causation that $E = e$ is an attempt to cater for ordinary linguistic practices, which is out of place in analyses of causal metaphysics, as we have previously suggested (Twardy and Korb, 2004).

D The Voting Machine

The Voting Machine example (Halpern and Pearl, 2004, Example A.3) showed that Hitchcock's redundancy range account (H2) was too strict in requiring that ϕ be unchanged. Consequently, Halpern & Pearl replaced that requirement with what we have called here resetting ranges.

Their voting scenario has two people, V_1 and V_2, vote, and the measure passes (P) if at least one votes in favor. Votes are counted by a voting machine M, so that the votes affect passage only via M. Actually, both vote in favor, and the measure passes. It is said that both votes should count as causes.



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

Figure 4: Voting Machine: V_1 = Vote 1, V_2 = Vote 2, M = Voting Machine, and P = Measure Passes.

To review Halpern & Pearl's analysis, consider V_1. In the actual background $V_2 = 1$, and V_1 made no difference. Let $V_2 = 0$, which is clearly in the redundancy range. Now $V_1 = 0$ results in $P = 0$, but $V_1 = 1$ results in $P = 1$ even if (only if!) M is reset to its actual state, $M = 1$. So $V_1 = 1$ is an actual cause, and by symmetry, so is $V_2 = 1$.

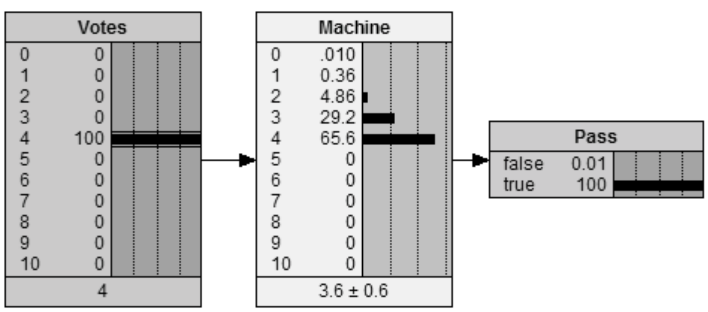
We can match their result, but we think the solution is not entirely satisfactory. On this analysis, $V_1 = 1$ would be an actual cause no matter how many other votes were cast: setting all other votes to 0 is still in the redundancy range. But as the number of votes N goes to infinity, the conclusion becomes increasingly implausible.

As before, we might seek refuge in probabilities, but to no avail. Suppose that all votes are in favor, but that each vote has a 10% chance of being tallied as against the measure. Whether that 10% is the chance of voting against the measure or being incorrectly tallied should not matter. For the case with two votes, consider V_1:

Background	Δ_1	Δ_2
Actual: $V_2 = 1$	.09	0
Redundant: $V_2 = 0$	.9	.09

V_1 is an actual cause for all reasonable values of ε. If $\varepsilon < .09$, then V_1 is an actual cause in the actual background, because it has a small Δ_1 that is nevertheless greater than ε. If ε is greater than that, but less than 0.9, then V_1 is an actual cause in a redundant background, where it has a large effect. We think that $\varepsilon > 0.9$ is not reasonable.

The general case with N positive votes can be treated as a binomial experiment with N trials, each with probability of success p. Therefore the mean and variance of M are $\mu = Np$, and $\sigma^2 = Np(1-p)$. Let $q = 1 - p$, in our case, 0.1. The probability the measure will fail is q^N, and the difference made by 1 vote is roughly q^{N-1}, which becomes vanishingly small. The distribution for 4 votes is shown here:



For $\varepsilon = 0$, of course each vote is an actual cause, but with a contribution rapidly approaching 0.

For all numbers of positive votes N, for moderate $\varepsilon < p$, each vote is still an actual cause. This is because we can always find a redundant background of k opposing votes where $\Delta_2 < \varepsilon$, but where the loss of one more vote (V_1) makes $\Delta_1 > \varepsilon$. This is because the counterfactual change from N to $N' = N - k$

positive votes will be an order of magnitude smaller than the subsequent change to $N' - 1$ positive votes.

However, for a fixed ε, as N gets larger, the plausibility of N' becomes smaller. In the case where $N = 10$, the counterfactual $N' = 4$ has only a .014% chance (.00014 probability) of occurring.

It would seem that the binary judgment of "actual cause" has to be supplemented with a notion of causal strength (see, e.g., Korb et al., 2011).

References

Joseph Y. Halpern and Judea Pearl. Causes and explanations: A Structural-Model approach - part i: Causes. *British Journal for the Philosophy of Science*, 2004. URL `http://www.cs.cornell.edu/home/halpern/abstract.html#actcaus`.

K. Korb, L. Hope, A. Nicholson, and K. Axnick. Varieties of causal intervention. In *PRICAI 2004: Trends in Artificial Intelligence*, page 322–331, 2004. URL `http://www.csse.monash.edu.au/~korb/pubs/intervene.pdf`.

Kevin B. Korb, Erik Nyberg, and Lucas R. Hope. A new causal power theory. In P.M.K. Illari, F. Russo, and J. Williamson, editors, *Causality in the Sciences*. Oxford Univ Pr, 2011.

Charles R. Twardy and Kevin B. Korb. A criterion of probabilistic causality. *Philosophy of Science*, in press, 2004.

Charles R. Twardy and Kevin B. Korb. Actual caustion by probabilistic active paths. *Philosophy of Science*, 78:???–???, December 2011.