Actual Causation

Holger Andreas and Mario Günther*

Forthcoming in *dialectica*

Abstract

We put forth an analysis of actual causation. The analysis centers on the notion of a causal model that provides only partial information as to which events occur. The basic idea is this: *c* causes *e* only if there is a causal model that is uninformative on *e* and in which *e* will occur if *c* does. We show that our analysis captures more causal scenarios than any account that tests for counterfactual dependence under certain contingencies.

We analyse causation between token events. Here is the gist of the analysis: an event *c* is a cause of another event *e* only if both events occur, and—after taking out the information whether or not *e* occurs—*e* will occur if *c* does. We will show that the analysis successfully captures a wide range of causal scenarios, including overdetermination, preemption, switches, and scenarios of double prevention. This set of scenarios troubles counterfactual accounts of actual causation. Even sophisticated counterfactual accounts still fail to deal with all of its members. And they fail for a principled reason: to solve overdetermination and preemption, they rely on a strategy which gives the wrong results for switches and a scenario of double prevention. Our analysis, by contrast, is not susceptible to this principled problem.

Counterfactual accounts try to analyse actual causation in terms of counterfactual dependence. An event *e* counterfactually depends on an event

*Mario.Guenther@lmu.de. The authors contributed equally.

c if and only if (iff), were c not to occur, e would not occur. Among the accounts in the tradition of Lewis (1973), counterfactual dependence between two occurring events is taken to be sufficient for causation.[1] That is, an occurring event c is a cause of a distinct occurring event e if, were c not to occur, e would not occur. Counterfactual accounts thus ask 'what would happen if the putative cause were absent?' Under this counterfactual assumption they claim causation if the presumed effect is absent as well.

Overdetermination is troublesome for counterfactual accounts. Consider the scenario depicted in Figure 1.

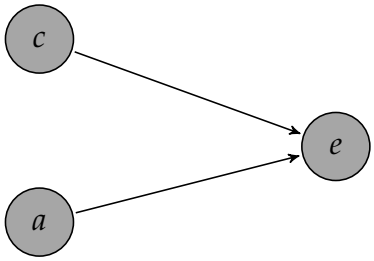


Figure 1

Neuron c and neuron a fire. The firing of each of c and a alone suffices to excite neuron e. Hence, the common firing of c and a overdetermines e to fire. Arguably, the firing of c is a cause of e's excitation, and so is the firing of a.

What would have happened had c not fired? If c had not fired, e would have been excited anyways. After all, a would still have fired. Hence, as is well known, c is not a cause of e on Lewis's (1973) account. More sophisticated accounts solve the scenario of overdetermination as follows: c's excitation is a cause of e's firing because e's firing counterfactually depends on c's excitation if a were not to fire. The non-actual contingency that a

[1] See Lewis (1973, 2000), Ramachandran (1997), Hitchcock (2001), Yablo (2002), Woodward (2003), Hall (2004, 2007), Halpern and Pearl (2005), Halpern (2015), and many others.

does not fire reveals a hidden counterfactual dependence of the effect e on its cause c. The general strategy is to test for counterfactual dependence under certain contingencies, be they actual or non-actual. We call counterfactual accounts relying on this strategy 'sophisticated'.[2]

Numerous sophisticated accounts analyse causation relative to a causal model. A causal model represents a causal scenario by specifying which events occur and how certain events depend on others. Formally, a causal model $\langle M, V \rangle$ is given by a variable assignment V and a set M of structural equations. For the above scenario of overdetermination, V may be given by the set $\{c, a, e\}$, which says that all neurons fire. M is given by $\{e = c \vee a\}$, which says that e fires iff c or a does. In this causal model, we may set the variable c to $\neg c$, a to $\neg a$ and propagate forward the changes effected by these interventions. Given that $\neg c$ and $\neg a$, the structural equation determines that $\neg e$. The equation tells us that e would not have fired, if c had not fired under the contingency that a had not fired. Hence, the above solution of overdetermination can be adopted: c is a cause of e (relative to the causal model) because e counterfactually depends on c if $\neg a$ is set by intervention.[3]

We solve the problem of overdetermination in a different way. The idea is this: remove enough information about which events occur so that there is no information on whether or not a putative effect occurs; an event c is then a cause of this effect only if—after the removal of information—the effect will occur if c does.

We use causal models to implement the idea. The result of the information removal is given by a causal model $\langle M, V' \rangle$ that provides only partial information as to which events occur, but complete information about the dependences between the events. To outline the preliminary analysis: c is a cause of e relative to a causal model $\langle M, V \rangle$ iff

[2]Sophisticated counterfactual accounts are, for example, provided by Ramachandran (1997), Hitchcock (2001), Yablo (2002), Woodward (2003, Ch. 2.7), Halpern and Pearl (2005), Hall (2007), and Halpern (2015).

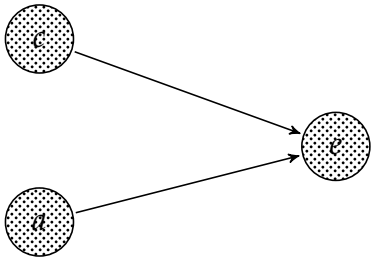
[3]Sophisticated accounts that rely on causal models are, for example, provided by Hitchcock (2001), Woodward (2003, Ch. 2.7), Halpern and Pearl (2005), Hall (2007), and Halpern (2015).

(1) c and e are true in $\langle M, V\rangle$, and

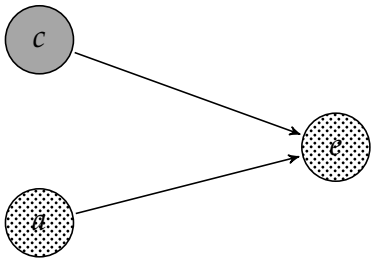
(2) there is $V' \subset V$ such that $\langle M, V'\rangle$ contains no information as to whether e is true, but in which e will become true if c does.

By these conditions, we test whether an event brings about another event in a causal scenario. Causation is here actual production.

Why is c's excitation a cause of e's firing in the overdetermination scenario? Take the causal model $\langle M, V'\rangle$ that contains no information about whether or not the effect e occurs:



Here, a neuron is dotted iff V' contains no information as to whether the neuron fires or not. Since all neurons are dotted, the causal model contains no information on which neurons fire. But it still contains all the information about dependences among the neurons, as encoded by the structural equation of the overdetermination scenario. Let us now intervene such that c becomes excited:



The structural equation is triggered and determines e to fire. Hence, c's excitation is a cause of e's firing on our analysis. The overdetermination scenario is solved without counterfactually assuming the absence of the cause and without invoking any contingency.

It should be noted that the recent counterfactual theories of Gallow (2021) and Andreas and Günther (2021) are not sophisticated in our sense: they do not test for counterfactual dependence under certain contingencies. And so they are not susceptible to the principled problem. Indeed, both theories solve the set of scenarios that troubles sophisticated accounts. The analysis of Andreas and Günther (2021) relies on a removal of information just like the analysis proposed here, and can thus be seen as its counterfactual counterpart. We will briefly and favourably compare our analysis to its counterfactual counterpart in the Conclusion.

In what follows, we refine our analysis, apply it to causal scenarios, and compare it to counterfactual accounts. In Section 1, we introduce our account of causal models. In Section 2, we state a preliminary version of our analysis and explain its rationale. We apply this analysis to various causal scenarios in Section 3. In response to certain switching scenarios, we amend our preliminary analysis by a condition of weak difference making. In Section 4, we state the final version of our analysis. In Section 5, we compare our analysis to the extant counterfactual accounts. Section 6 concludes the paper.

1 Causal Models

In this section, we explain the basic concepts of causal models. Our account parallels the account of causal models in Halpern (2000). Unlike Halpern, we introduce structural equations as formulas and not as functions. Another difference is that our account is confined to binary variables, the values of which are represented by literals.[4] We will see shortly that these modelling choices allow us to define causal models in a straight-

[4]With a few modifications, both the framework and the analysis can be extended to non-binary variables.

forward way, in particular causal models that carry only partial information as to which events occur. In the appendix, we supplement the explanations of the core concepts of causal models with precise definitions.

Our causal models have two components: a set M of structural equations and a consistent set V of literals. Where p is a propositional variable, p is a positive literal and $\neg p$ a negative literal. We give literals a semantic role. The literals in V denote which events occur and which do not, that is, which events and absences are actual. $p \in V$ means that the event corresponding to p occurs. $\neg p \in V$, by contrast, means that no token event p of the relevant type occurs. Since the set of literals is consistent, it cannot be that both p and $\neg p$ are in V. Arguably, an event cannot both occur and not occur at the same time.

A structural equation denotes whether an event would occur if some other events were or were not to occur. Where p is a propositional variable and ϕ a propositional formula, we say that

$$p = \phi$$

is a structural equation. Each logical symbol of ϕ is either a negation, a disjunction, or a conjunction. ϕ can be seen as a truth function whose arguments represent occurrences and non-occurrences of events. The truth value of ϕ determines whether p or $\neg p$.

Consider the scenario of overdetermination depicted in Figure 1. There are arrows from the neurons c and a to the neuron e. The arrows represent that the propositional variable e is determined by the propositional variables c and a. The specific structural equation of the overdetermination scenario is $e = c \vee a$. This equation says that e occurs iff c or a does. A set of structural equations describes dependences between actual and possible token events.

For readability, we will represent causal models in two-layered boxes. The causal model of the overdetermination scenario, for example, is given by $\langle \{e = c \vee a\}, \{c, a, e\} \rangle$. We will depict such causal models $\langle M, V \rangle$ in a box, where the upper layer shows the set M of structural equations and the lower layer the set V of actual literals. For the overdetermination scenario, we obtain:

$e = c \vee a$
c, a, e

We say that a set V of literals satisfies a structural equation $p = \phi$ just in case both sides of the equation have the same truth value when plugging in the literals in V. In the case of overdetermination, the actual set of literals satisfies the structural equation. By contrast, the set of literals $\{c, a, \neg e\}$ does not satisfy $e = c \vee a$. When plugging in the literals, the truth values of e and $c \vee a$ do not match. We say that a set V of literals satisfies a set M iff V satisfies each member of M.

The structural equations and the literals determine which events occur and which do not occur in a causal model. This determination can be expressed by a relation of satisfaction between a causal model and a propositional formula.

Definition 1. $\langle M, V \rangle$ **satisfies** ϕ
$\langle M, V \rangle$ satisfies ϕ iff ϕ is true in all complete sets V^c of literals that extend V and satisfy M. A set V^c of literals is complete iff each propositional variable (in the language of M) is assigned to a truth value by V^c.

If V is complete, this definition boils down to: $\langle M, V \rangle$ satisfies ϕ iff V satisfies ϕ, or V does not satisfy M. Provided V is complete, $\langle M, V \rangle$ satisfies at least one of ϕ and $\neg\phi$ for any formula ϕ.

Our analysis relies on causal models that contain no information as to whether or not an effect occurs. We say that a causal model $\langle M, V \rangle$ is *uninformative* about a formula ϕ iff $\langle M, V \rangle$ satisfies none of ϕ and $\neg\phi$. Note that $\langle M, V \rangle$ cannot be uninformative on any formula if V is complete.

In the scenario of overdetermination, the causal model $\langle M, V \rangle$ is uninformative on e for $V = \emptyset$. There are four complete extensions that satisfy $M = \{e = c \vee a\}$. One of these is $\{\neg c, \neg a, \neg e\}$. Hence, $\langle M, V \rangle$ does not satisfy e. Similarly, $\langle M, V \rangle$ does not satisfy $\neg e$. There is a complete extension of V that satisfies M but fails to satisfy $\neg e$. The actual set $\{c, a, e\}$ of literals, for example, but also the sets $\{c, \neg a, e\}$ and $\{\neg c, a, e\}$. The structural equation constrains the overdetermination scenario to four possible

cases. These cases are expressed by the complete sets of literals which satisfy M.

Why is $\langle M, V\rangle$ not uninformative on e for $V = \{a\}$? Well, there is no complete extension of V that satisfies the structural equation in M but fails to satisfy e. There are only two such complete extensions: $\{c, a, e\}$ and $\{\neg c, a, e\}$. If a remains in the set V of literals, e is determined independent of whether or not c occurs.

It remains to introduce interventions. Recall that a structural equation $p = \phi$ determines the truth value of the variable p if certain variables q occurring in ϕ are given truth values by the literals in V. To represent an intervention that sets p to one of the truth values, we replace the equation $p = \phi$ by the corresponding literal p or $\neg p$. We implement such interventions by the notion of a submodel. M_I is a submodel of M relative to a consistent set I of literals just in case M_I contains the literals in I and the structural equations of M for the variables which do not occur in I. In symbols,

$$M_I = \{(p = \phi) \in M \mid p \notin I \text{ and } \neg p \notin I\} \cup I.$$

We denote interventions by an operator $[\cdot]$ that takes a model M and a consistent set of literals I, and returns a submodel. In symbols, $M[I] = M_I$. In the overdetermination scenario, for instance, we may intervene on $M = \{e = c \vee a\}$ by $\{\neg a\}$. This yields: $M[\{\neg a\}] = \{\neg a, e = c \vee a\}$. The causal model $\langle M_{\{\neg a\}}, \emptyset\rangle$ satisfies $\neg a$, and $\langle M_{\{\neg a\}}[\{\neg c\}], \emptyset\rangle$ satisfies $\neg e$. If $\neg c$ were actual under the contingency that $\neg a$, $\neg e$ would be actual.

Finally, note that the above definition of satisfaction applies to causal models and causal submodels. The definition does not only capture the relation of a causal model $\langle M, V\rangle$ satisfying a formula ϕ, but also the relation of a causal submodel $\langle M_I, V\rangle$ satisfying such a formula. This is explained further in the appendix.

2 The Analysis

We are now in a position to spell out our analysis in a more precise way. The key idea is as follows: for c to be a cause of e, there must be a causal

model $\langle M, V' \rangle$ that is uninformative about e, while intervening by c determines e to be true. The latter condition must be preserved under all interventions by a set A of actual events. In more formal terms:

Definition 2. Actual Cause (Preliminary)
Let $\langle M, V \rangle$ be a causal model such that V satisfies M. c is an actual cause of e relative to $\langle M, V \rangle$ iff

(C1) $\langle M, V \rangle$ satisfies c and e, and

(C2) there is $V' \subset V$ such that $\langle M, V' \rangle$ is uninformative on e, while for all $A \subseteq V$, $\langle M_A[\{c\}], V' \rangle$ satisfies e.

The rationale behind our analysis is straightforward: there must be a way in which a genuine cause actually brings about its effect. This production of the effect can be reconstructed by means of a causal model $\langle M, V' \rangle$ that contains some information of the original causal model $\langle M, V \rangle$, but no information about whether the effect is actual. Or so requires condition (C2).

Furthermore, (C2) says production of an effect must respect actuality. The idea is that the causal process initiated by a genuine cause must respect what actually happened. A genuine cause cannot produce its effect via non-actual events and absences. The process from cause to effect must come about as it actually happened. This idea requires that a genuine cause must bring about its effect by events and absences that are actual. We implemented this requirement as follows: intervening upon the uninformative model $\langle M, V' \rangle$ by any subset of the actual events and absences V must preserve that e will become actual if c does. Thereby, it is ensured that a genuine cause cannot bring about its effect by events or absences that are not actual. If c is a genuine cause, there can be no subset A of the actual literals V that interferes with the determination of e by c in the respective uninformative model. We describe this feature of (C2) as *intervention by actuality*.

3 Scenarios

In this section, we test our analysis of actual causation against causal scenarios, and compare the results to the counterfactual accounts due to Lewis (1973), Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015). We follow Paul and Hall (2013, p. 10) in laying out the structure of causal scenarios by neuron diagrams. "Neuron diagrams earn their keep", they write, "by representing a complex situation clearly and forcefully, allowing the reader to take in at a glance its central causal characteristics."[5] We introduce simple neuron diagrams for which there is always a corresponding causal model. Our causal models, however, can capture more causal scenarios than simple neuron diagrams.

A neuron diagram is a graph-like representation that comes with different types of arrows and different types of nodes. Any node stands for a neuron, which fires or else does not. The firing of a neuron is visualized by a gray-shaded node, the non-firing by a white node. For the scenarios to be considered, we need two types of arrows. Each arrow with a head represents a stimulatory connection between two neurons, each arrow ending with a black dot an inhibitory connection. Furthermore, we distinguish between *normal* neurons that become excited if stimulated by another and *stubborn* neurons whose excitation requires two stimulations. Normal neurons are visualized by circles, stubborn neurons by thicker circles. A neuron diagram obeys four rules. First, the temporal order of events is left to right. Second, a normal neuron will fire if it is stimulated by at least one and inhibited by none. Third, a stubborn neuron will fire if it is stimulated by at least two and inhibited by none. Fourth, a neuron will not fire if it is inhibited by at least one.

Typically, neuron diagrams are used to represent events and absences. The firing of a neuron indicates the occurrence of some event and the non-firing indicates its non-occurrence. Recall that we analyse causation between token events relative to a causal model $\langle M, V \rangle$, where the causal model represents the causal scenario under consideration. We thus need a

[5] This being quoted, there are some shortcomings of neuron diagrams. For details, see Hitchcock (2007b).

correspondence between neuron diagrams and causal models.

Here is a recipe to translate an arbitrary neuron diagram, as detailed here, into a causal model. Given a neuron diagram, the corresponding causal model can be constructed in a step-wise fashion:

For each neuron n of the neuron diagram,

(i) assign n a propositional variable p.

(ii) If n fires, add the positive literal p to the set V of literals.

(iii) If n does not fire, add the negative literal $\neg p$ to V.

(iv) If n has an incoming arrow, write on the right-hand side of p's structural equation a propositional formula ϕ such that ϕ is true iff n fires.[6]

This recipe adds a positive literal p to the set V of literals for each neuron that fires, and a negative literal $\neg p$ for each neuron that dos not fire. Then the neuron rules are translated into structural equations. One can thus read off a neuron diagram its corresponding causal model: if a neuron is shaded gray, p is in the set V of literals of the corresponding causal model; if a neuron is white, $\neg p$ is in V.

We have already added a feature to neuron diagrams in the introduction. Recall that dotted nodes represent neurons about which there is no information as to whether or not they fire. In more formal terms, if $p \notin V$ and $\neg p \notin V$, the corresponding neuron will be dotted. We portray now how our analysis solves the problems posed by overdetermination, conjunctive

[6]The structural equations can be explicitly constructed from the rules governing neuron diagrams. That is, the catch-all condition (iv) can be replaced by the following clauses. (v) For each stimulatory arrow ending in a normal neuron n, add disjunctively to the right side of p's structural equation the variable that corresponds to the neuron where the arrow originates. (vi) For each pair of stimulatory arrows ending in a stubborn neuron n, add disjunctively to the right side of p's structural equation the conjunction of the two variables that correspond to the two neurons where the arrows originate. (vii) For each inhibitory arrow ending in n, add conjunctively to the right side of p's structural equation the negation of the variable that corresponds to the neuron where the arrow originates. This translation shows that there is a principled transition from simple neuron diagrams to our causal models.

causes, early and late preemption, switches, prevention, and two scenarios of double prevention.

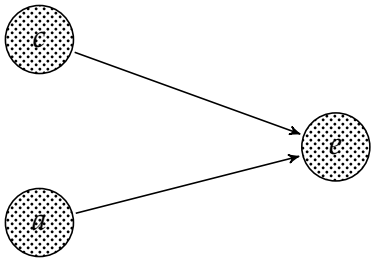
3.1 Overdetermination

Scenarios of overdetermination are commonly represented by the neuron diagram depicted in Figure 1. Here is a story that fits the structure of overdetermination: A prisoner is shot by two soldiers at the same time (c and a), and each of the bullets is fatal without any temporal precedence. Arguably, both shots should qualify as causes of the death of the prisoner (e).

Our recipe translates the neuron diagram of Figure 1 into the following causal model $\langle M, V \rangle$:

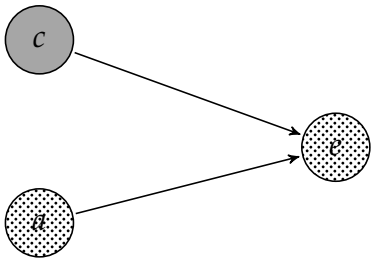
$e = c \vee a$
c, a, e

Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$e = c \vee a$

Intervening by $\{c\}$ yields:



c
$e = c \vee a$

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e. And intervening by any subset of actual events does not undo the determination.[7] In more detail, any intervention by a subset of $\{c, a, e\}$ yields a causal model that determines e to be true. Due to the symmetry of the scenario, a is a cause of e.[8]

Overdetermination is trouble for the counterfactual account of Lewis (1973). There, Lewis defines actual causation as the transitive closure of counterfactual dependence between occurring events. Let c and e be distinct events. c is a cause of e iff c and e occur, and there is a sequence $\langle c, d_1, ..., d_n, e \rangle$ of distinct events and absences such that each element in the sequence (except the first) counterfactually depends on its predecessor in a non-backtracking way.[9] Recall that e counterfactually depends on c just in case if c were not to occur, e would not occur. Lewis insists that each counterfactual in the series of counterfactual dependences is non-backtracking.[10] A backtracking counterfactual retraces some past causes from an effect: if the effect e were not to occur, its past causes c and a must have been absent. Intuitively, this backtracking counterfactual is true in the confines of the overdetermination scenario. Yet Lewis does not allow such backtracking counterfactuals to figure in the series of counterfactual dependences.

It follows from Lewis's account that non-backtracking counterfactual dependence between occurring events is sufficient for causation. As soon as c and e occur, there is a sequence $\langle c, e \rangle$. If, in addition, e counterfactually

[7] We will not always explicitly mention this intervention by actuality in the scenarios to come.

[8] The final analysis of Section 4 counts the set $\{c, a\}$ as a cause of e.

[9] Lewis (1986, p. 189) says that an absence $\neg a$ is the non-occurrence of any event of type A. If the absence $\neg a$ had not been, some token event a of type A would have been. Counterfactual dependence between occurring events is thus only a special case of counterfactual dependence between actual events and absences. The latter is still sufficient for causation, or so argues Lewis.

[10] See Lewis (1986, p. 201), Lewis (1973, p. 566), and Lewis (1979, p. 456-9). Lewis (1979, p. 456) characterises reasoning by backtracking as follows: "We know that present conditions have their past causes. [...] if the present were different then these past causes would have to be different". The exclusion of backtracking counterfactuals plays a crucial role in Lewis's analysis of causation. For subtle details regarding backtracking counterfactuals see Lewis (1979).

depends on c in a non-backtracking way, c is a cause of e. In the scenario of overdetermination, c is not a cause of e on this account.[11] There is no suitable series of counterfactual dependences. If c had not fired, e would have been excited all the same. After all, a would still have fired and excited e. Due to the symmetry of the scenario, a is not a cause of e either. But then, what caused the death of the prisoner? Surely, we do not want to say that the death is uncaused.

The counterfactual accounts of causation due to Hitchcock (2001) and Halpern and Pearl (2005) solve the scenario of overdetermination as follows: c is a cause of e because e counterfactually depends on c if $\neg a$ is set by intervention. Their tests for causation allow for non-actual contingencies, that is, to set variables to non-actual values and to keep them fixed at these non-actual values. We will see that this feature is problematic in switching scenarios and extended double prevention.

Halpern (2015) modifies the Halpern and Pearl (2005) definition of actual causation. The main difference is that the modified definition admits only actual contingencies for the counterfactual test. Hence, the modified definition fails to recognize the individual overdeterminers as actual causes, while it counts the set $\{c, a\}$ of overdeterminers to be an actual cause of e.[12] It has troubles to handle overdetermination, as already pointed out by Andreas and Günther (2021). This indicates that overdetermination haunts counterfactual accounts to date.

3.2 Conjunctive Causes

In a scenario of conjunctive causes, an effect occurs only if two causes obtain. The following neuron diagram depicts a scenario of conjunctive causes:

[11] Lewis (2000) refines his earlier account. There, the idea to hold certain events fixed while altering others surfaces (p. 191). However, he does not advocate to keep certain merely possible events or absences fixed. Hence, his refined account does not solve overdetermination either.

[12] This being said, Halpern (2015) calls each conjunct of an actual cause *part of a cause*.

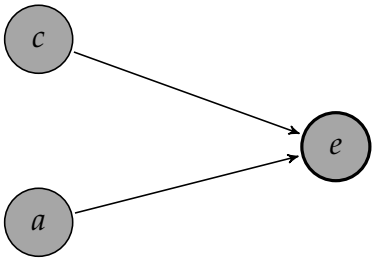


Figure 2

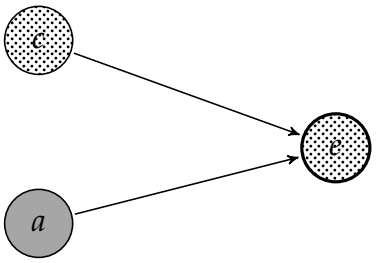
The neurons c and a fire. Together they bring the stubborn neuron e to fire. Had one of c and a not fired, e would not have been excited. Hence, the firing of both neurons is necessary for e's excitation.

Our recipe translates the neuron diagram of Figure 2 into the following causal model $\langle M, V \rangle$:

$e = c \wedge a$
c, a, e

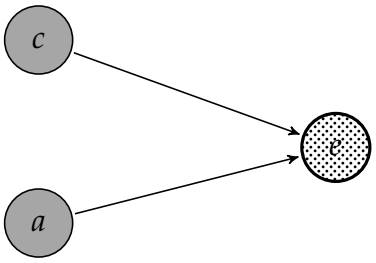
The scenario of conjunctive causes differs from the scenario of overdetermination only in the structural equation for e. While the structural equation is *disjunctive* in the scenario of overdetermination, here the equation is *conjunctive*. The occurrence of both events, c and a, is necessary for e to occur.

Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$e = c \wedge a$
a

Intervening by $\{c\}$ yields:



c
$e = c \wedge a$
a

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e. Again, due to the symmetry of the scenario, a is a cause of e.[13]

At first sight, conjunctive causes seem to be no problem for counterfactual accounts. If c had not fired, e would not have fired. Hence, on the counterfactual accounts, c is a cause of e. And by the symmetry of the scenario, a is a cause of e. However, the accounts due to Lewis (1973) and Hitchcock (2001) do not allow sets of events to be causes, unlike the definitions of actual causation provided by Halpern and Pearl (2005) and Halpern (2015). Yet the latter definitions still do not count the set containing *c and a* as an actual cause of e in this scenario of *conjunctive* causes. Hence, none of these counterfactual accounts counts the set containing the two individual causes as a cause of the effect. This is peculiar for reasons worked out by Andreas and Günther (2021).

3.3 Early Preemption

Preemption scenarios are about backup processes: there is an event c that, intuitively, causes e. But even if c had not occurred, there is a backup event a that would have brought about e. Paul and Hall (2013, p. 75) take the following neuron diagram as canonical example of early preemption:

[13] The final analysis of Section 4 counts the set $\{c, a\}$ as a cause of e.

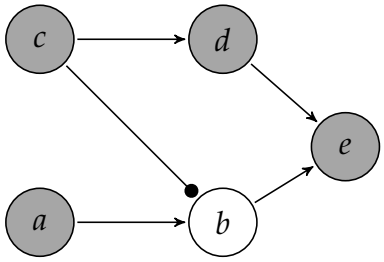


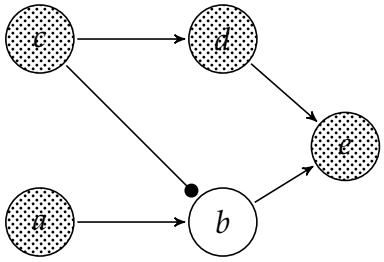
Figure 3

c's firing excites neuron d, which in turn leads to an excitation of neuron e. At the same time, c's firing inhibits the excitation of b. Had c not fired, however, a would have excited b, which in turn would have led to an excitation of e. The actual cause c preempts the mere potential cause a.[14]

Our recipe translates the neuron diagram of early preemption into the following causal model $\langle M, V \rangle$:

$d = c$
$b = a \wedge \neg c$
$e = d \vee b$
$c, a, d, \neg b, e$

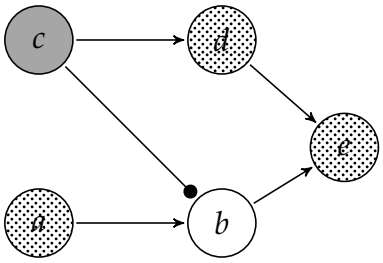
Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$d = c$
$b = a \wedge \neg c$
$e = d \vee b$
$\neg b$

[14]Following Halpern and Pearl (2005, pp. 861-2), we take the model of symmetric overdetermination in Section 3.1 to be inappropriate for representing the structure of the early preemption scenario.

Intervening by $\{c\}$ yields:



c
$d = c$
$b = a \wedge \neg c$
$e = d \vee b$
$\neg b$

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e.

Relative to $\langle M, V \rangle$, a is not a cause of e. The reason is that actuality intervenes. The causal model $\langle M, V' \rangle$ is uninformative on e only for $V' = \emptyset$ or $V' = \{\neg b\}$. Intervening on $\langle M, V' \rangle$ by $\{\neg b\}$ yields a causal model in which a does not produce e, independently of the choice of V'. In more formal terms, $\langle M_{\{\neg b\}}[\{a\}], V' \rangle$ does not satisfy e. For each choice of V', there is a complete extension that satisfies the structural equations $a, \neg b, d = c$, and $e = d \vee b$ but does not satisfy e. This extension of V' is $\{a, \neg b, \neg c, \neg d, \neg e\}$. Intuitively, a is not a genuine cause of e since a would produce e only via an event b that did not actually occur. Hence, a is not a cause of e because a does not *actually* produce e.

Lewis's (1973) account solves early preemption. In Figure 3, c is a cause of e. Both occur and there is a sequence $\langle c, d, e \rangle$ such that e counterfactually depends in a non-backtracking way on d, and d does so on c. The counterfactual 'if d had not fired, its cause c would have to have not fired' is backtracking. Barring backtracking, we do not obtain that b would have fired because c did not and thus b would not be inhibited. Hence, if d had not fired, b would still not have fired. And so 'If d had not fired, e would not have fired' comes out true under the non-backtracking requirement. a is not a cause of e. For there is no sequence of events and absences from a to e where each counterfactually depends on its predecessor in a non-backtracking way. If b had fired, e would still have fired.

The solution to early preemption by Hitchcock (2001) and Halpern and Pearl (2005) is analogous to their solution for overdetermination. c is a

cause of e because e counterfactually depends on c under the contingency that $\neg b$. By contrast to their solution for overdetermination, the contingency is actual in cases of early preemption. Hence, Halpern's (2015) account solves early preemption as well.

3.4 Late Preemption

Lewis (1986, p. 200) subdivides preemption into early and late. We have discussed early preemption in the previous section: a backup process is cut off before the process started by the preempting cause brings about the effect. In scenarios of late preemption, by contrast, the backup process is cut off only because the genuine cause brings about the effect before the preempted cause could do so. Lewis (2000, p. 184) provides the following story for late preemption:

> Billy and Suzy throw rocks at a bottle. Suzy throws first, or maybe she throws harder. Her rock arrives first. The bottle shatters. When Billy's rock gets to where the bottle used to be, there is nothing there but flying shards of glass. Without Suzy's throw, the impact of Billy's rock on the intact bottle would have been one of the final steps in the causal chain from Billy's throw to the shattering of the bottle. But, thanks to Suzy's preempting throw, that impact never happens.

Crucially, the backup process initiated by Billy's throw is cut off only by Suzy's rock impacting the bottle. Until her rock impacts the bottle, there is always a backup process that would bring about the shattering of the bottle an instant later.[15]

Halpern and Pearl (2005, pp. 861-2) propose a causal model for late preemption, which corresponds to the following neuron diagram:

[15]The problem posed by late preemption can be solved by fine-grained individuation conditions for events. According to these conditions, the shattering of the bottle and the shattering of the bottle an instant later are two different events. By adopting this strategy counterfactual accounts run into the trouble of spurious causation: they identify causal relations where, intuitively, there are none. See, for instance, Lewis (1986, p. 204-5), Collins et al. (2004, pp. 45-8) and Paul and Hall (2013, Ch. 3.4.2).

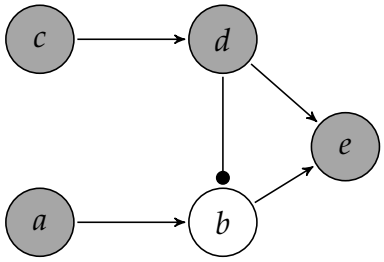


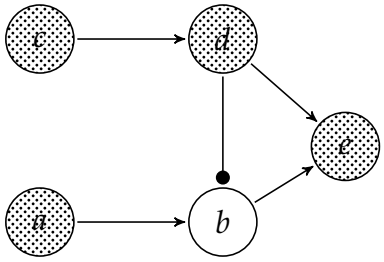
Figure 4

Suzy throws her rock (c) and Billy his (a). Suzy's rock impacts the bottle (d), and so the bottle shatters (e). Suzy's rock impacting the bottle (d) prevents Billy's rock from impacting the bottle ($\neg b$). (The "inhibitory signal" from d takes "no time" to arrive at b.)

Our recipe translates the neuron diagram of late preemption into the following causal model $\langle M, V \rangle$:

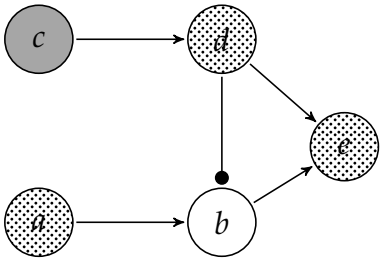
$d = c$
$b = a \wedge \neg d$
$e = d \vee b$
$c, a, d, \neg b, e$

Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$d = c$
$b = a \wedge \neg d$
$e = d \vee b$
$\neg b$

Intervening by $\{c\}$ yields:



c
$d = c$
$b = a \wedge \neg d$
$e = d \vee b$
$\neg b$

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e.

Relative to $\langle M, V \rangle$, a is not a cause of e. The intuitive reason is that Billy's rock did not actually impact the bottle. The formal reasoning is perfectly analogous to the one for the scenario of early preemption in the previous section. Our analysis solves early and late preemption in a uniform manner.

Lewis's (1973) account does not solve late preemption. Suzy's throw (c) is not a cause of the bottle shattering (e). There is no sequence $\langle c, ..., e \rangle$ of events and absences such that each event (except c) counterfactually depends on its predecessor in a non-backtracking way. There is, of course, the sequence $\langle c, d, e \rangle$, and if Suzy had not thrown ($\neg c$), her rock would not have impacted the bottle ($\neg d$). However, if Suzy's rock had not impacted the bottle ($\neg d$), the bottle would have shattered anyways (e). The reason is that—on a non-backtracking reading—if Suzy's rock had not impacted the bottle ($\neg d$), Billy's rock would have (b). But if Billy's rock had impacted the bottle (b), it would have shattered (e). By contrast to scenarios of early preemption, there is no chain of stepwise dependences that run from cause to effect: there is no sequence of non-backtracking counterfactual dependences that links Suzy's throw and the bottle's shattering.[16]

The counterfactual accounts of causation due to Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) solve the scenario of late preemption analogous to early preemption. c is a cause of e because e counterfactually depends on c under the contingency that $\neg b$.

[16]Lewis (2000) claims to solve late preemption. This claim is highly controversial. See, for instance, Paul (1998).

3.5 Simple Switch

In switching scenarios, some event f helps to determine the causal path by which some event e is brought about (Hall, 2000, p. 205). The following neuron diagram represents a simple version of a switching scenario:

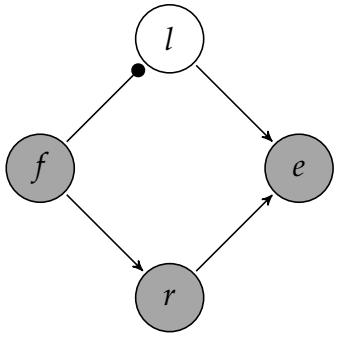


Figure 5

The firing of neuron f excites r's firing, which in turn excites neuron e. At the same time, f's firing inhibits the excitation of l. The neuron l is a little special: it would have been excited in case f had not fired. f determines which one of l and r is firing, and thus determines the causal path by which e is excited. We say f acts like a switch as to e.

Let us supplement our neuron diagram by a story due to Hall (2007, p. 28). Flipper is standing by a switch in the railroad tracks. A train approaches in the distance. She flips the switch (f), so that the train travels down the right track (r), instead of the left (l). Since the tracks reconverge up ahead, the train arrives at its destination all the same (e). We agree with Hall that flipping the switch is not a cause of the train's arrival. The story assumes that flipping the switch makes no difference to the train's arrival: 'the train arrives at its destination all the same'. The flipping merely switches the causal path by which the train arrives.[17]

[17]There is a noteworthy difference between switching scenarios and scenarios of preemption. If the non-actual switch position $\neg f$ were actual, $\neg f$ would help bring about e. By contrast, if it were actual that the genuine cause does not occur in scenarios of preemption, it's absence would not help bring about the effect. If Suzy were not to throw her rock, her not throwing would not help to bring about the bottle's shattering.

Our recipe translates the neuron diagram of the switching scenario into the following causal model $\langle M, V \rangle$:

$l = \neg f$ $r = f$ $e = l \vee r$
$f, \neg l, r, e$

Relative to $\langle M, V \rangle$, f is not a cause of e. The reason is that there exists no causal model $\langle M, V' \rangle$ uninformative on e. Any complete extension of the empty set V' of literals that satisfies the structural equations of M contains e. In fact, there are only two complete extensions that satisfy the structural equations, viz. the actual $\{f, \neg l, r, e\}$ and the non-actual $\{\neg f, l, \neg r, e\}$. The structural equations in M determine e no matter what.[18]

Our analysis requires for c to be a cause of e that there must be a causal model uninformative about e in which c brings about e. The idea is that, for an event to be caused, it must arguably be possible that the event does not occur. However, in the switching scenario, there is no causal model uninformative on e in the first place. Hence, f is not a cause of e in the simple switch.

A simplistic counterfactual analysis says that an event c is a cause of a distinct event e just in case both events occur, and e would not occur if c had not occurred. This suggests that the switching scenario is no challenge for counterfactual accounts, because e would occur even if f had not. And yet it turns out that cases like the switching scenario continue to be troublesome for counterfactual accounts.

Recall that Lewis (1973) defines actual causation to be the transitive closure of non-backtracking counterfactual dependence between occurring events. In the switching scenario, f, r, and e occur, and both r counterfactually depends on f in a non-backtracking way and e does so on r. Barring

[18] Hall (2007, p. 118) writes that the 'basic' switch in Paul and Hall (2013, p. 232) has "the obvious causal model": $M = \{b = a, l = b \wedge f, r = b \wedge \neg f, e = l \vee r\}, V = \{a, b, f, l, \neg r, e\}$. Relative to this causal model, our analysis says that f is not a cause of e, as desired. Relative to the causal scenario, where the equation for e is replaced by $e = l$, our analysis says that f is a cause of e, as desired (Paul and Hall, 2013, p. 235).

backtracking, if r had not fired, e would not have fired. By the transitive closure imposed on the one-step causal dependences, Lewis (1973) is forced to say that f is a cause of e.[19]

The sufficiency of (non-backtracking) counterfactual dependence for causation is widely shared among the accounts in the tradition of Lewis, for instance by Hitchcock (2001), Woodward (2003), Hall (2004, 2007), and Halpern and Pearl (2005). However, the counterfactual accounts based on structural equations reject the transitivity of causation. Still, Hitchcock (2001) counts f to be a cause of e. The reason is that there is an active causal path from f over r to e and keeping the off-path variable l fixed at its actual value induces a counterfactual dependence of e on f. Similarly, Halpern and Pearl (2005) and Halpern (2015) count f to be a cause of e, since e counterfactually depends on f under the actual contingency that $\neg l$. Hence, even the contemporary counterfactual accounts misclassify f to be a cause of e.[20] Allowing for actual contingencies solved preemption, but leads to trouble in switching scenarios. Without allowing for actual contingencies, it is unclear how the counterfactual accounts solve preemption. It seems as if the sophisticated counterfactual accounts have no choice here but to take one hit.

3.6 Realistic Switch

The representation of switching scenarios is controversial. Some authors criticize the simple switch in Figure 5 from the previous section because they believe that any 'real-world' event has more than one causal influence (e.g. Hitchcock (2009, p. 396)). The idea is that the train can only pass on the right track because nothing blocks the track, it is in good conditions, and so on. These critics insist on 'realistic' scenarios in which there is always more than just one event that causally affects another. The simple switch is thus inappropriate because there must be another neuron whose

[19] Lewis (2000, pp. 194-5) still imposes transitivity on his refined analysis of causation. As a result, the refined anlysis is also forced to say that f is a cause of e in the switching scenario.

[20] Halpern (2015) uses normality considerations to solve the present switching scenario. See Blanchard and Schaffer (2017) for a criticism of this strategy.

firing is necessary for the excitation of l. Some authors then quickly point out that the causal model of the resulting switch is indistinguishable from the one of early preemption (e.g. Beckers and Vennekens (2018, pp. 848-51)). And this is a problem for any account of causation that only relies on causal models. For c should intuitively be a cause of e in early preemption, but f should not be a cause in a 'realistic' switching scenario.[21]

It is too quick to point out that switches and early preemption are structurally indistinguishable. After all, the critics who insist on 'realistic' scenarios are bound to say that there should also be another neuron whose firing is necessary for the excitation of r. This restores the symmetry between l and r which seems to be essential to switching scenarios. The following neuron diagram depicts our realistic switch:

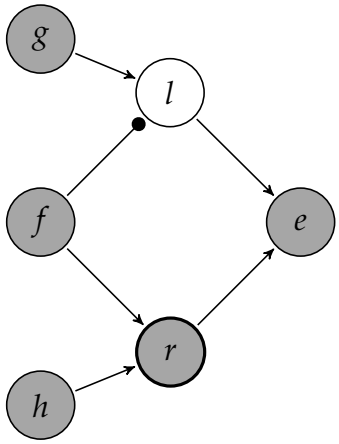


Figure 6

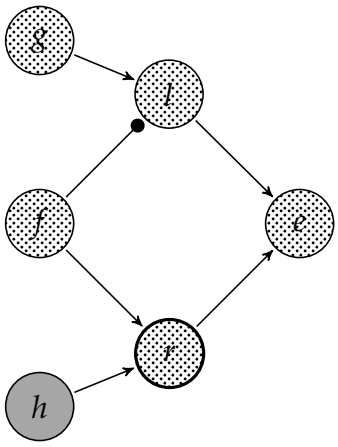
The joint firing of neurons f and h excites r's firing, which in turn excites neuron e. At the same time, f's firing inhibits the excitation of l. Had f not fired, the firing of g would have excited l, which in turn would have excited e. In the actual circumstances, f determines which one of l and r is firing, and thus acts like a switch as to e.

[21]The problem posed by structurally indistinguishable causal models where our intuitive causal judgments differ is further discussed in Section 5.1.

Our recipe translates the neuron diagram of our realistic switch into the following causal model $\langle M, V \rangle$:

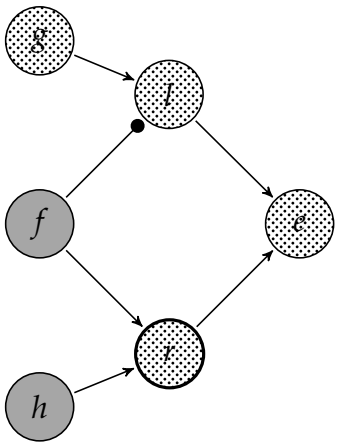
$l = g \wedge \neg f$ $r = f \wedge h$ $e = l \vee r$
$g, f, h, \neg l, r, e$

Relative to $\langle M, V \rangle$, f is a cause of e according to our preliminary analysis. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$l = g \wedge \neg f$ $r = f \wedge h$ $e = l \vee r$
h

Intervening by $\{f\}$ yields:



f $l = g \wedge \neg f$ $r = f \wedge h$ $e = l \vee r$
h

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{f\}}, V' \rangle$ satisfies e. Our preliminary analysis wrongly counts the 'realistic switch' f as a cause of e.

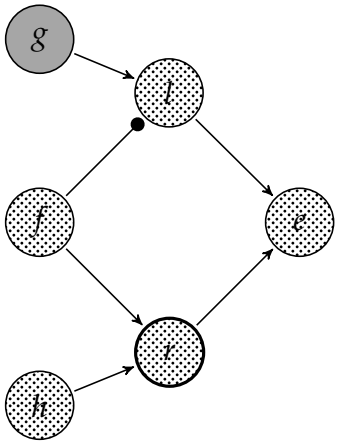
It is time to amend our preliminary analysis by a condition of *weak difference making*. The idea is this: if some event c is a cause of an event e, then it is not the case that $\neg c$ would be a cause of the same event e. Sartorio (2006, p. 75) convinces us that this principle of weak difference making is a condition "the true analysis of causation (if there is such a thing) would have to meet".[22] But this condition is violated by 'realistic switches': f helps to bring about an effect e, and so would the non-actual $\neg f$. So a 'realistic switch' is not a cause if we demand of any genuine cause c of some effect e that $\neg c$ would not also bring about e. We demand that $\neg c$ would not also bring about e by the following condition:

(C3) There is no $V'' \subset V \setminus \{c\}$ such that $\langle M, V'' \rangle$ is uninformative on e and $\langle M[\{\neg c\}], V'' \rangle$ satisfies e.

(C3) demands that there is no causal model uninformative on e in which e is actual if $\neg c$ is. The condition ensures that a cause is a difference maker in the weak sense that its presence and its absence could not bring about the same effect. This implies Sartorio's principle of weak difference making: if c is a cause of e, then $\neg c$ would not also be a cause of e. And note that our condition of difference making is weaker than the difference-making requirement of (sophisticated) counterfactual accounts of causation. Unlike them, we do not require that $\neg e$ is actual under the supposition that $\neg c$ is actual (given certain contingencies).

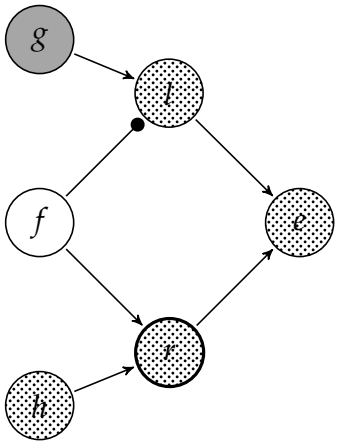
(C3) ensures that f is not a cause of e in the realistic switch. For this to be seen, consider the following causal model $\langle M, V'' \rangle$ that is uninformative on e.

[22] For more details, see Andreas and Günther (2020, pp. 1584&1590).



$l = g \wedge \neg f$ $r = f \wedge h$ $e = l \vee r$
g

Intervening by $\{\neg f\}$ yields:



f $l = g \wedge \neg f$ $r = f \wedge h$ $e = l \vee r$
h

Obviously, this causal model determines e to be true. In more formal terms, $\langle M_{\{f\}}, V'' \rangle$ satisfies e. Our preliminary analysis amended by (C3) says that the 'realistic switch' f is not a cause of e, as desired.[23] We will

[23]Hitchcock (2009, pp. 395-6) modifies Paul and Hall's (2013) 'basic' switch of fn. 18. The modified switch has the 'obvious causal model': $M = \{b = a, l = g \wedge b \wedge f, r = b \wedge h \wedge \neg f, e = l \vee r\}, V = \{a, g, b, h, f, l, \neg r, e\}$. Relative to this causal model, (C3) rules out that f is a cause of e, as desired. Halpern and Hitchcock (2010, p. 16) and Halpern (2016, p. 72) propose to model the train scenario by the following causal model: $M = \{e = (f \wedge \neg lb) \vee (\neg f \wedge \neg rb)\}, V = \{f, \neg lb, \neg rb, e\}$. The variables rb and lb indicate whether or not the right and left track are blocked, respectively. Relative to this causal model, (C3)

leave it as an exercise for the reader to check that (C3) does not undo any causes our preliminary definition identifies in this paper, except for the 'realistic switches'.

Lewis's (1973) account misclassifies f as a cause of e in our realistic switch. As in the simple switch, there is a causal chain running from f to e: the sequence $\langle f, r, e\rangle$ of actual events such that each event (except f) counterfactually depends on its predecessor in a non-backtracking way. Similarly, Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) all misclassify f as a cause of e. The reasons are analogous to the reasons in the simple switch. Roughly, e counterfactually depends on f when l is fixed at its actual value.

3.7 Prevention

To prepare ourselves for a discussion of double prevention, let us take a look at simple prevention first. Paul and Hall (2013, p. 174) represent the basic scenario of prevention by the following neuron diagram:

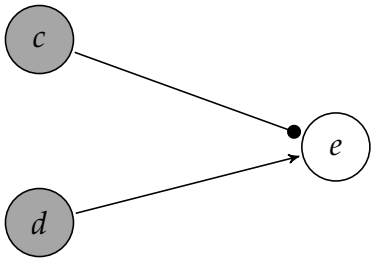


Figure 7

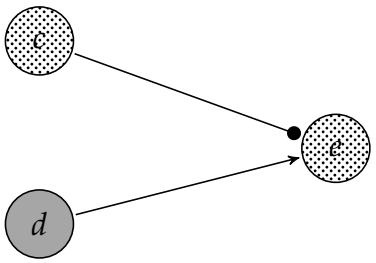
Neuron c fires and thereby inhibits that neuron e gets excited. e would have been excited by d if the inhibitory signal from c were absent. But as it is, c prevents e from firing. That is, c causes $\neg e$ by prevention.

Our recipe translates the neuron diagram of prevention into the following causal model $\langle M, V\rangle$:

rules out that f is a cause of e, as desired.

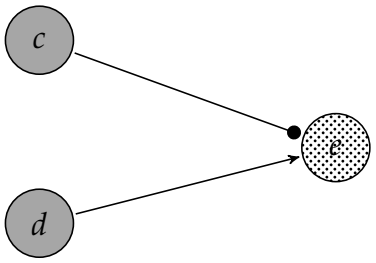
$e = \neg c \wedge d$
$c, d, \neg e$

Relative to $\langle M, V \rangle$, c is a cause of $\neg e$. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on $\neg e$.



$e = \neg c \wedge d$
d

Intervening by $\{c\}$ yields:



c
$e = \neg c \wedge d$
d

Obviously, this causal model determines $\neg e$ to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies $\neg e$. Moreover, d is not a cause of $\neg e$ relative to $\langle M, V \rangle$. Any causal model $\langle M, V' \rangle$ uninformative on $\neg e$ must be uninformative on c as well. Intervening by d in $\langle M, V' \rangle$ does not determine $\neg e$.

Counterfactual accounts face no challenge here. If c had not fired, e would have fired. Counterfactual dependence between actual events and absences is sufficient for causation. Hence, c is a cause of $\neg e$. If d had not fired, e would not have fired, even under the contingency that c did not fire. Hence, d is not a cause of $\neg e$.

3.8 Double Prevention

Double prevention can be characterized as follows. c is said to double prevent e if c prevents an event that, had it occurred, would have prevented e. In other words, c double prevents e if c cancels a threat for e's occurrence. Paul and Hall (2013, pp. 154, 175) represent an example of double prevention by the following neuron diagram:

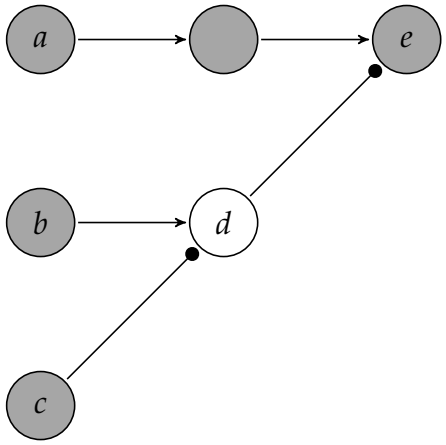


Figure 8

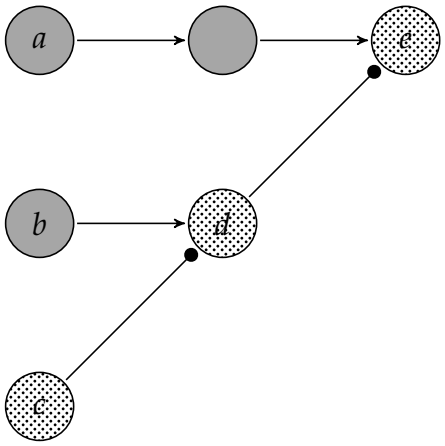
c's firing prevents d's firing, which would have prevented e's firing. The example of double prevention exhibits a counterfactual dependence: given that b fires, e's firing counterfactually depends on c's firing. If c did not fire, d would fire, and thereby prevent e from firing. Hence, c's firing double prevents e's firing in Figure 8. In other words, c's firing cancels a threat for e's firing, viz. the threat originating from b's firing.

Paul and Hall (2013) say that c is a cause of e in the scenario of Figure 8. They thereby confirm that there is causation by double prevention. e counterfactually depends on c. Hence, the accounts of causation due to Lewis (1973, 2000), Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) agree with Paul and Hall in counting c a cause of e. How does our account fare?

Our recipe translates the neuron diagram of double prevention into the following causal model $\langle M, V \rangle$:

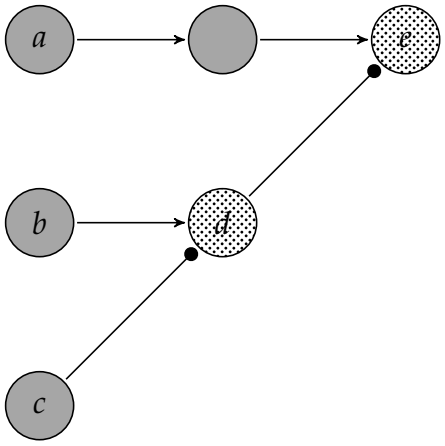
$d = b \wedge \neg c$ $e = a \wedge \neg d$
$a, b, c, \neg d, e$

Relative to $\langle M, V \rangle$, c is a cause of e. For this to be seen, consider the following causal model $\langle M, V' \rangle$ that is uninformative on e.



$d = b \wedge \neg c$ $e = a \wedge \neg d$
a, b

Intervening by $\{c\}$ yields:



c $d = b \wedge \neg c$ $e = a \wedge \neg d$
a, b

Obviously, this causal model determines $\neg d$ and so e to be true. In more formal terms, $\langle M_{\{c\}}, V' \rangle$ satisfies e.

3.9 Extended Double Prevention

Hall (2004, p. 247) presents an extension of the scenario depicted in Figure 8. The extended double prevention scenario fits the structure of the following neuron diagram:

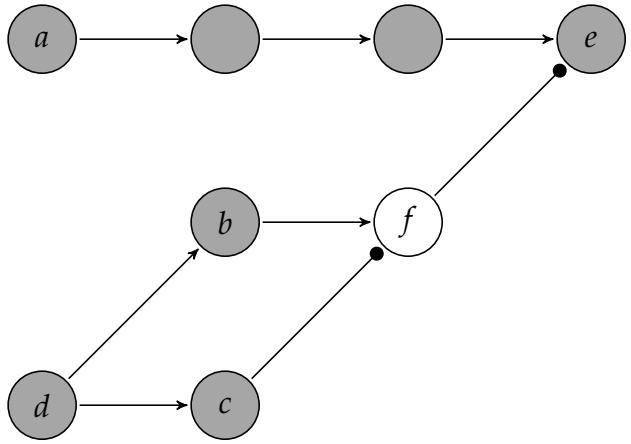


Figure 9

Figure 9 extends Figure 8 by neuron d, which figures as a common cause of b and c. d starts a process via b that threatens to prevent e. At the same time, d initiates another process via c that prevents the threat. d cancels its *own* threat—the threat via b—to prevent e. In the example of the previous section, the threat originated independent of its preventer. Here, by contrast, d creates and cancels the threat to prevent e. This difference is sufficient for d not to be a cause of e, or so argue for instance Paul and Hall (2013, p. 216). Observe that the structure characteristic of double prevention is embedded in Figure 9. The firing of neuron c inhibits f's firing that, had it fired, would have inhibited e's firing. Nevertheless, this scenario of double prevention exhibits an important difference to its relative of the previous section: e does not counterfactually depend on d. If d had not fired, e would still have fired.

Hitchcock (2001, p. 276) provides a story that matches the structure of the scenario. A hiker is on a beautiful hike (a). A boulder is dislodged (d) and rolls toward the hiker (b). The hiker sees the boulder coming and ducks

(c), so that he does not get hit by the boulder ($\neg f$). If the hiker had not ducked, the boulder would have hit him, in which case the hiker would not have continued the hike. Since, however, he was clever enough to duck, the hiker continues the hike (e).

Hall (2007, p. 36) calls the subgraph $d - b - c - f$ a *short circuit* with respect to e: the boulder threatens to prevent the continuation of the hike, but provokes an action that prevents this threat from being effective. Like switching scenarios, the scenario seems to show that there are cases where causation is not transitive: the dislodged boulder d produces the ducking of the hiker c, which in turn enables the hiker to continue the hike e. But it is counterintuitive to say that the dislodging of the boulder d causes the continuation of the hike e. After all, the dislodgement of the boulder is similar to a switch as to the hiker *not* getting hit by the boulder: d helps to bring about $\neg f$, and if $\neg d$ were actual, $\neg d$ would also help to bring about $\neg f$. In this sense, d is causally inert.

Our recipe translates the neuron diagram of the boulder scenario into the following causal model $\langle M, V \rangle$:

$b = d$
$c = d$
$f = b \wedge \neg c$
$e = a \wedge \neg f$
$a, d, b, c, \neg f, e$

Relative to $\langle M, V \rangle$, d is not a cause of e. The reason is that the causal model $\langle M, V' \rangle$ is only uninformative on e if a is not in V'. But then $\langle M_{\{d\}}, V' \rangle$ does not satisfy e.

In words, the causal model $\langle M, V' \rangle$ is uninformative about e only if a is not in the set V' of literals. But then intervening with d does not make e true. After all, a is necessary for determining e. If we were to keep a in the literals, the model would not be uninformative. There is no complete extension of $V' = \{a\}$ that satisfies all the structural equations of M but fails to satisfy e.

On Lewis's (1973) account, d is a cause of e. There is a sequence $\langle d, c, \neg f, e\rangle$ of events and absences such that each element (except d) counterfactually depends on its predecessor in a non-backtracking way. The structural equation accounts of Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) classify d as a cause of e. The reason is that e counterfactually depends on d under the contingency that b.

The situation is bad for the sophisticated counterfactual accounts. While their general strategy to allow for possibly non-actual contingencies solves overdetermination and preemption, it is the very same strategy that is at fault for the unintuitive results in the switching scenario and extended double prevention. The backfiring of their general strategy casts doubt on whether it was well motivated in the first place. If the general strategy is merely motivated by solving overdetermination, it turns out that overdetermination still haunts the sophisticated accounts of causation. By contrast to these counterfactual accounts, our analysis of actual causation solves overdetermination without further ado. Our analysis has thus a major advantage over the sophisticated counterfactual accounts.

4 Final Analysis

In Section 1, we stated a preliminary version of our analysis and amended it in Section 3.6 by condition (C3). The amended version is still preliminary because it assumes that both the cause and the effect are single events. This assumption is violated in certain causal scenarios. Recall, for instance, the scenario of conjunctive causes from Section 3.2. There, two events are necessary for an effect to occur, and so the set containing the two events should count as a cause of said effect. To give an example, lightning resulted in a forest fire only because of a preceding drought. Here, it seems plausible that lightning together with the preceding drought is an—if not *the*—cause of the forest fire.[24]

We lift the restriction of cause and effect to single literals as follows. A

[24] Andreas and Günther (2021, pp. 608-10) argue that it is desirable if an account of causation can count sets of events as causes.

cause is a set of literals C, an effect an arbitrary Boolean formula. Where C is a set of literals, $\bigwedge C$ stands for the conjunction of all literals in C and $\neg C$ for the negation of all literals in C. Our final analysis of actual causation can now be stated.

Definition 3. Actual Cause
Let $\langle M, V\rangle$ be a causal model such that V satisfies M. C is a set of literals and ε a formula. C is an actual cause of ε relative to $\langle M, V\rangle$ iff

(C1*) $\langle M, V\rangle$ satisfies $\bigwedge C \wedge \varepsilon$, and

(C2*) there is $V' \subset V$ such that $\langle M, V'\rangle$ is uninformative on ε, while for all $A \subseteq V$ and all non-empty $C' \subseteq C$, $\langle M_A[C'], V'\rangle$ satisfies ε; and

(C3*) there is no $V'' \subset V \setminus C$ such that $\langle M, V''\rangle$ is uninformative on ε and $\langle M[\neg C], V''\rangle$ satisfies ε.

In this more general analysis, clause (C2*) contains a minimality condition ensuring that any cause contains only causally relevant literals. For this to be seen, suppose there is a set $C' \subset C$ whose members are causally irrelevant for ε. That is, intervening by C' in any partial model uninformative on ε does not make ε true (under all interventions by actuality). Then, by the minimality condition, C would not be a cause, contrary to our assumption. Thanks to this condition, causally irrelevant factors cannot simply be added to genuine causes.[25]

How fare the counterfactual accounts with respect to sets of causes? Let us consider the scenario of overdetermination. As explained in Section 4.1, Halpern's (2015) account counts only the set of individual causes as a genuine cause. The other counterfactual accounts do not count this set as a cause. We think it is reasonable to recognize both the individual causes and the set of these causes as a proper cause. We would say that, for instance, two soldiers shooting a prisoner, where each bullet is fatal without any temporal precedence, is a perfectly fine cause for the death of the prisoner. The shooting of the two soldiers brings about the death of the prisoner.

[25] If one wants cause and effect to be distinct, one should amend Definition 3 by a clause like this: no element of C occurs in ε.

The account of Hitchcock (2001) does not admit causes that are sets of variables. Hence, the set containing the two individual causes does not count as a cause in the scenarios of overdetermination and conjunctive causes. Unlike Hitchcock's account, the accounts due to Halpern and Pearl (2005) and Halpern (2015) admit causes to be sets of variables. Still, these accounts do not recognize the set containing the two individual causes as a cause in the scenario of conjunctive causes. The accounts share the same minimality condition according to which a strict superset of a cause cannot be a cause. Hence, they are forced to say that, for instance, the drought together with the lightning is not a cause of the forest fire *because* one of these events (and indeed both) already counts as a cause for this effect. This reason for why the set is not a cause is a little odd.

5 Comparison

In this section, we compare our analysis to the considered counterfactual accounts. First, we focus on the results of the different accounts. Then we compare—on a conceptual level—our analysis to the counterfactual accounts that rely on causal models.

5.1 Results

The results of our analysis and of the considered counterfactual accounts are summarized in the following table. We abbreviate the accounts of Lewis (1973), Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015) by $\mathcal{L}$'73, Hitch'01, HP'05, and H'15, respectively.

Causes of e or $\neg e$	$\mathcal{L}$'73	Hitch'01	HP'05	H'15	Author(s)
Overdetermination	–	c, a	c, a	$\{c, a\}$	$c, a, \{c, a\}$
Conjunctive Causes	c, a	c, a	c, a	c, a	$c, a, \{c, a\}$
Early Preemption	c	c	c	c	c
Late Preemption	–	c	c	c	c
Switches	f	f	f	f	–
Prevention	c	c	c	c	c
Double Prevention	c	c	c	c	c
E. Double Prevention	d	d	d	d	–

None of the counterfactual accounts listed in the table provides the intuitively correct results for the simple and 'realistic' switching scenarios and extended double prevention. Lewis's (1973) account misclassifies f and d as causes of e, respectively, because of the transitive closure he imposes on the step-wise and non-backtracking counterfactual dependences. And without imposing transitivity, his analysis of causation cannot solve early preemption. For Halpern (2015), Hitchcock (2001) and Halpern and Pearl (2005), the reason for the misclassification is that they allow for actual contingencies. And if they were not to allow for such, their accounts would fail to solve preemption. The counterfactual accounts due to Hitchcock (2001) and Halpern and Pearl (2005) solve overdetermination, but only by allowing for even non-actual contingencies.

We have thus shown that the sophisticated counterfactual accounts fail to capture the set of overdetermination, preemption, switches, and extended double prevention. And they fail for a principled reason: they can solve overdetermination and preemption only if they allow for contingencies. But, by allowing for contingencies, they fail to solve the switching scenario and extended double prevention. If they were not to allow for contingencies, they would solve the switching scenario and extended double prevention, but it would be unclear how they could solve overdetermination and preemption. Our analysis, by contrast, does not fall prey to such a principled problem.

Let us summarize the verdicts about the results, where ✓, ✗, and **!** stand for correct, false, and partially correct, respectively.

Causes of e or $\neg e$	$\mathcal{L}$'73	Hitch'01	HP'05	H'15	Author(s)
Overdetermination	✗	✓	✓	!	✓
Conjunctive Causes	!	!	!	!	✓
Early Preemption	✓	✓	✓	✓	✓
Late Preemption	✗	✓	✓	✓	✓
Switch	✗	✗	✗	✗	✓
Prevention	✓	✓	✓	✓	✓
Double Prevention	✓	✓	✓	✓	✓
E. Double Prevention	✗	✗	✗	✗	✓

There remains another problem to be solved. The problem concerns any account that relies on simple causal models which only factor in structural equations and values of variables (or our sets of literals). Such accounts face pairs of scenarios for which our causal judgments differ, but which are structurally indistinguishable. Overdetermination, for instance, is isomorphic to bogus prevention. In bogus prevention, an event p would prevent another event d. But, as it is, there is no event c present that would bring about d in the first place. Hence, the preventer p and the absence of c overdetermine that d does not occur. By contrast to overdetermination, however, the preventer p is intuitively not a cause of the absence $\neg d$. Since the accounts of Hitchcock (2001) and Halpern and Pearl (2005) consider only structural equations and the values of variables, they cannot distinguish between p and one of the causes in overdetermination. The former must be falsely classified to be a cause if the latter is correctly classified so.[26] And our analysis has the same problem.[27]

Hitchcock (2007a), Hall (2007), Halpern (2008), Halpern and Hitchcock

[26] As pointed out by Hiddleston (2005, p. 32) and Hall (2007, p. 44), Hitchcock's (2001) and Halpern and Pearl's (2005) allowance of non-actual contingencies solves the overdetermination scenario, but it leads to the intuitively wrong results in *bogus* cases of both prevention and double prevention. From this perspective, the non-actual contingencies, as opposed to merely actual contingencies, are thus even more bad news.

[27] This being said, the causal model of bogus prevention is: $M = \{d = \neg p \wedge c\}, V = \{\neg c, p, \neg d\}$. Blanchard and Schaffer (2017, p. 200-2) argue that this causal model is inappropriate for bogus prevention and propose to model the bogus scenario by a model isomorphic to early preemption. If they are right, our analysis would give the correct verdict for bogus prevention. We would like to thank an anonymous referee for this observation.

(2015), and Halpern (2015) all aim to solve the problem of isomorphism by taking into account default or normality considerations. This additional factor gives considerable leeway to solve some of the isomorphic pairs. However, actual causation does not seem to be default-relative, as pointed out by Blanchard and Schaffer (2017). They also show that the accounts amended by a notion of default still face counterexamples and even invite new ones. Nevertheless, the problem of isomorphism suggests that simple causal models ignore a factor that impacts our intuitive causal judgments. We think this ignored factor are not default considerations, but a meaningful distinction between events that occur and events that do not. After all, a distinction between events and absences seems to be part of the structure of causation. Yet current accounts relying on causal models are blind to such a distinction.

Our analysis of causation is thus incomplete. We need to amend it by a meaningful distinction between events and absences, which allows us to tackle the problem of isomorphism. More generally, we miss an account of what constitutes an appropriate causal model. That is, an account that tells us which causal models are appropriate for a given causal scenario. For now, we have just assumed that the causal models obtained from simple neuron diagrams are appropriate. This assumption already smuggled in certain metaphysical assumptions about events. We will elaborate these underpinnings of our analysis elsewhere.

5.2 Conceptual Differences

Let us compare—on a more conceptual level—our analysis to the counterfactual accounts that likewise rely on causal models. As we have seen, these sophisticated counterfactual accounts analyse actual causation in terms of contingent counterfactual dependence relative to a causal model. Hitchcock (2001), Halpern and Pearl (2005), and Halpern (2015), for instance, have put forth such accounts. All of these accounts have in common that the respective causal model provides full information about what actually happens, and what would happen if the state of affairs were different. Hence, causal models allow them to test for counterfactual dependence: provided c and e are actual in a causal model, would $\neg e$ be

actual if $\neg c$ were? If so, e counterfactually depends on c; if not, not.

The mentioned accounts put forth more elaborate notions of counterfactual dependence. These notions specify which variables other than c and e are to be kept fixed by intervention when testing for counterfactual dependence. The accounts ask a test question for contingent counterfactual dependence: relative to a causal model, where c and e are actual, would $\neg e$ be actual if $\neg c$ were under the contingency that certain other variables are kept fixed at certain values? If so, e counterfactually depends on c under the contingency; if not, not. To figure out whether c is a cause of e, counterfactual accounts propagate forward—possibly under certain contingencies—the effects of the counterfactual assumption that a putative cause were absent.

We analyse, by contrast, actual causation in terms of production relative to a causal model that provides only partial information. More specifically, our analysis relies on models that carry no information with respect to a presumed effect e: they are uninformative as to whether or not the event or absence e is actual. Such uninformative models allow us to test whether an actual event or absence is actually produced by another. The test question goes as follows: in a model uninformative on e, will e become actual if c does? If so, c is a producer of e; if not, not. And a producer c is then a cause of e if $\neg c$ would not also be a producer of e.

Our test has no need that $\neg e$ becomes actual if $\neg c$ were actual. Instead the question is whether, in an uninformative model, an actual event produces (and makes a weak difference to) another in accordance with what actually happened. The novelty of our account is not so much to consider actual production, but to consider production in a causal model that is uninformative on the presumed effect. As a consequence, when testing for causation, we never intervene on a causal model, where the set of actual literals is complete. This stands in stark contrast to counterfactual accounts which always intervene on causal models, where each variable is assigned a value.

On our analysis, c is a cause of e only if c produces e under *all* interventions by actuality. There is a mentionable symmetry to Halpern's (2015) account which allows only for actual contingencies. On this account, c is a cause

of e if *there is* an intervention by actuality such that the actual e counterfactually depends on the actual c.[28] Production under all interventions by actuality is *necessary* for causation on our account, whereas counterfactual dependence between actual events under some intervention by actuality is *sufficient* on Halpern's.

Counterfactual notions of causation generally say that a cause is necessary for an effect: without the cause, no effect. By contrast, our notion of causation says that a cause is sufficient for its effect given certain background conditions. The background conditions are given by the partial set of literals of the causal model that is uninformative on the effect. That is, these conditions are jointly not sufficient for the effect given the structural equations. However, together with a genuine cause these conditions are jointly sufficient for the effect (given the same structural equations). Relative to the causal model uninformative on the effect, a cause is thus necessary and sufficient for its effect.[29]

6 Conclusion

We have put forth an analysis of actual causation. In essence, c is a cause of e just in case c and e are actual, and there is a causal model uninformative on e in which c actually produces e, and there is no such uninformative causal model in which $\neg c$ would produce e. Our analysis successfully captures various causal scenarios, including overdetermination, preemption,

[28] The intervention by actuality on Halpern's (2015) account can just be the intervention by the empty set.

[29] Perhaps, our analysis bears more resemblance to regularity analyses of causation than to counterfactual accounts. The core idea behind regularity analyses can be glossed as follows: c is a cause of e just in case, given the laws of nature, c together with a minimal set of background conditions is jointly sufficient for e. Indeed, our analysis of causation can be seen as a regularity theory when one replaces 'laws of nature' by 'structural equations' and 'minimal set of background conditions' by 'partial set of actual literals'. In a causal model uninformative on e, intervening by a cause c is sufficient to bring about the effect e. In a very specific sense, this says that the 'laws' and 'minimal background conditions' imply that c is sufficient for e. However, we are not aware of any regularity theory that employs an equivalent to our uninformative models.

switches, and extended double prevention. All extant sophisticated counterfactual accounts of causation fail to capture at least two of the causal scenarios considered. With respect to this set, our analysis is strictly more comprehensive than those accounts.

The sophisticated counterfactual accounts, which rely on causal models, run into problems for a principled reason. They fail to solve the switching scenario and extended double prevention because they allow for possibly non-actual contingencies when testing for counterfactual dependence. Such contingencies are needed to solve the problems of overdetermination and preemption. Our analysis, by contrast, is neither premised on counterfactuals of the form "if $\neg c$, then $\neg e$", nor on considering possibly non-actual contingencies. Hence, our analysis escapes the principled problem to which the sophisticated counterfactual accounts are susceptible.

The present analysis of causation has a counterfactual counterpart due to Andreas and Günther (2021). The counterfactual analysis likewise relies on an information removal and uninformative causal models. The gist is this: an event c is a cause of another event e just in case both events occur, and—after removing the information whether or not c and e occur—e would not occur if c were not to occur. This analysis does not rely on the strategy common to the sophisticated counterfactual accounts, and is therefore also not susceptible to their principled problem.

The two analyses largely come to the same verdicts. However, unlike the present preliminary analysis, the preliminary counterfactual analysis cannot identify the overdetermining causes in scenarios of symmetric overdetermination. And while the present final analysis counts the set $\{c, a\}$ as a cause in the scenario of conjunctive causes, the final counterfactual analysis does not. More importantly, the present final analysis does not count 'realistic switches' as causes, whereas the final counterfactual analysis does. The present analysis has therefore a slight edge over its counterfactual counterpart.

Acknowledgements. We would like to thank Frank Jackson, Philip Pettit, Katie Steele, Atoosa Kasirzadeh, Cei Maslen, Alan Hájek, Phil Dowe, and Daniel Stoljar for helpful comments on this paper. We are furthermore grateful to the anonymous reviewers for *dialectica*. We are happy for the

opportunities to present parts of this paper to the Philosophy Departmental Seminar at The Australian National University, the 2019 Annual Conference of the New Zealand Association of Philosophers, and the conference Bayesian Epistemology: Perspectives and Challenges at the Munich Center for Mathematical Philosophy.

Appendix: The Framework of Causal Models

In this appendix, we supplement the explanations of the core concepts of causal models with precise definitions. Let P be a set of propositional variables such that every member of P represents a distinct event. $\mathcal{L}_P$ is a propositional language that is defined recursively as follows: (i) Any $p \in P$ is a formula. (ii) If ϕ is a formula, then so is $\neg\phi$. (iii) If ϕ and ψ are formulas, then so are $\phi \vee \psi$ and $\phi \wedge \psi$. (iv) Nothing else is a formula.

As is well known, the semantics of a propositional language centers on the notion of a value assignment. A value assignment $v : P \mapsto \{T, F\}$ maps each propositional variable on a truth value. We can represent a value assignment, or valuation for short, in terms of literals. The set $L(v)$ yields the set of literals that represents the valuation v.

Definition 4. $L(v)$
Let $v : P \mapsto \{T, F\}$ be a valuation of the language $\mathcal{L}_P$. $L(v)$ is the set of literals of $\mathcal{L}_P$ such that, for any $p \in P$, (i) $p \in L(v)$ iff $v(p) = T$, and (ii) $\neg p \in L(v)$ iff $v(p) = F$.

We say that a set V of literals is complete—relative to $\mathcal{L}_P$—iff there is a valuation v such that $L(v) = V$. If the language is obvious from the context, we simply speak of a complete set of literals, leaving the parameter P implicit.

The function $L(v)$ defines a one-to-one correspondence between the valuations of $\mathcal{L}_P$ and the complete sets of $\mathcal{L}_P$ literals. In more formal terms, $L(v)$ defines a bijection between the set of valuations of $\mathcal{L}_P$ and the set of the complete sets of $\mathcal{L}_P$ literals. Hence, the inverse function $L^{-1}(V)$ of $L(V)$ is well defined for complete sets V of literals. Using the inverse of

$L(V)$, we can define what it is for a complete set V of literals to satisfy an $\mathcal{L}_P$ formula ϕ:

$$V \models \phi \text{ iff } L^{-1}(V) \models_C \phi, \qquad (V \models \phi)$$

where $\models_C$ stands for the satisfaction relation of classical propositional logic. In a similar vein, we define the semantics of a single structural equation:

$$V \models p = \phi \text{ iff, } L^{-1}(V) \models_C p \text{ iff } L^{-1}(V) \models_C \phi. \qquad (V \models p = \phi)$$

In simpler terms, V satisfies the structural equation $p = \phi$ iff both sides of the equation have the same truth value, on the valuation specified by V. We say that a set V of literals satisfies a set M of structural equations and literals iff V satisfies each member in M. In symbols,

$$V \models M \text{ iff } V \models \gamma \text{ for each } \gamma \in M. \qquad (V \models M)$$

These two relations of satisfaction in place, we can say what it is for a causal model $\langle M, V \rangle$ to satisfy a Boolean formula ϕ.

Definition 5. $\langle M, V \rangle \models \phi$
Let $\langle M, V \rangle$ be a causal model relative to $\mathcal{L}_P$. $\langle M, V \rangle \models \phi$ iff $V^c \models \phi$ for all complete sets V^c of literals such that $V \subseteq V^c$ and $V^c \models M$.

The definition says that ϕ is true in $\langle M, V \rangle$ iff it is true in all complete interpretations V^c that extend V and that satisfy M. For complete models, the definition boils down to $\langle M, V \rangle \models \phi$ iff $V \models \phi$ or $V \not\models M$.

There remains to define the notion of a submodel M_I that is obtained by an intervention I on a model M.

Definition 6. Submodel M_I
Let M be a set of structural equations of the language $\mathcal{L}_P$. Let I be a consistent set of literals. M_I is a submodel of M iff:

$$M_V = \{(p = \phi) \in M \mid p \notin I \text{ and } \neg p \notin I\} \cup I.$$

A submodel M_I has two types of members. First, the structural equations of M for those variables which do not occur in I. Second, the literals in

I. Hence, the syntactic form of a submodel M_I differs from the one of a model M. If I is non-empty, the submodel M_I has at least one member that is not a structural equation but a literal. The satisfaction relation $V \models M_I$ remains nonetheless well defined. The reason is that $V \models \gamma$ has been defined for both a structural equation γ and an $\mathcal{L}_P$ formula.

References

Andreas, Holger and Günther, Mario (2020). Causation in Terms of Production. *Philosophical Studies* **177**(6): 1565–1591.

Andreas, Holger and Günther, Mario (2021). Difference-Making Causation. *The Journal of Philosophy* **118**(12): 680–701.

Andreas, Holger and Günther, Mario (2021). A Ramsey Test Analysis of Causation for Causal Models. *The British Journal for the Philosophy of Science* **72**(2): 587–615.

Beckers, Sander and Vennekens, Joost (2018). A Principled Approach to Defining Actual Causation. *Synthese* **195**(2): 835–862.

Blanchard, Thomas and Schaffer, Jonathan (2017). Cause Without Default. In *Making a Difference*, edited by H. Beebee, C. Hitchcock, and H. Price, Oxford: Oxford University Press. pp. 175–214.

Collins, John , Hall, Ned , and Paul, Laurie (2004). Counterfactuals and Causation: History, Problems, and Prospects. In *Causation and Counterfactuals*, edited by J. Collins, N. Hall, and L. Paul, MIT Press.

Gallow, Dmitri J. (2021). A Model-Invariant Theory of Causation. *The Philosophical Review* **130**(1): 45–96.

Hall, Ned (2000). Causation and the Price of Transitivity. *The Journal of Philosophy* **97**(4): 198.

Hall, Ned (2004). Two Concepts of Causation. In *Causation and Counterfactuals*, edited by J. Collins, N. Hall, and L. Paul, MIT Press. pp. 225–276.

Hall, Ned (2007). Structural Equations and Causation. *Philosophical Studies* **132**(1): 109–136.

Halpern, Joseph (2016). *Actual Causality*. Cambridge, MA: MIT Press.

Halpern, Joseph Y. (2000). Axiomatizing Causal Reasoning. *Journal of Artificial Intelligence Research* **12**(1): 317–337.

Halpern, Joseph Y. (2008). Defaults and Normality in Causal Structures. In *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning*, edited by G. Brewka and J. Lang. Menlo Park, CA: AAAI Press, 198–208.

Halpern, Joseph Y. (2015). A Modification of the Halpern-Pearl Definition of Causality. *Proc. 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)* : 3022–3033.

Halpern, Joseph Y. and Hitchcock, Christopher (2010). Actual Causation and the Art of Modeling. In *Heuristics, Probability, and Causality: a Tribute to Judea Pearl*, edited by R. Dechter, H. Geffner, and J. Y. Halpern, London: College Publications. 383–406.

Halpern, Joseph Y. and Hitchcock, Christopher (2015). Graded Causation and Defaults. *British Journal for the Philosophy of Science* **66**(2): 413–457.

Halpern, Joseph Y. and Pearl, Judea (2005). Causes and Explanations: A Structural-Model Approach. Part I: Causes. *British Journal for the Philosophy of Science* **56**(4): 843–887.

Hiddleston, Eric (2005). Causal Powers. *The British Journal for the Philosophy of Science* **56**(1): 27–59.

Hitchcock, Christopher (2001). The Intransitivity of Causation Revealed in Equations and Graphs. *The Journal of Philosophy* **98**(6): 273–299.

Hitchcock, Christopher (2007a). Prevention, Preemption, and the Principle of Sufficient Reason. *Philosophical Review* **116**(4): 495–532.

Hitchcock, Christopher (2007b). What's Wrong with Neuron Diagrams? In *Causation and Explanation*, edited by J. K. Campbell, M. O'Rourke, and H. S. Silverstein, MIT Press. pp. 4–69.

Hitchcock, Christopher (2009). Structural equations and causation: six counterexamples. *Philosophical Studies* **144**(3): 391–401.

Lewis, David (1973). Causation. *The Journal of Philosophy* **70**(17): 556–567.

Lewis, David (1979). Counterfactual Dependence and Time's Arrow. *Noûs* **13**(4): 455–476.

Lewis, David (1986). Postscripts to "Causation". In *Philosophical Papers. Volume II*, edited by D. Lewis, Oxford University Press. pp. 172–213.

Lewis, David (2000). Causation as Influence. *The Journal of Philosophy* **97**(4): 182–197.

Paul, Laurie (1998). Problems with Late Preemption. *Analysis* **58**(1): 48–53.

Paul, Laurie and Hall, Ned (2013). *Causation: A User's Guide*. Oxford.

Ramachandran, Murali (1997). A Counterfactual Analysis of Causation. *Mind* **106**(422): 263–277.

Sartorio, Carolina (2006). Disjunctive Causes. *The Journal of Philosophy* **103**(10): 521–538.

Woodward, James (2003). *Making Things Happen : A Theory of Causal Explanation*. Oxford University Press.

Yablo, Stephen (2002). De Facto Dependence. *The Journal of Philosophy* **99**(3): 130–148.