A Lewisian Regularity Theory

Holger Andreas and Mario Günther*

Forthcoming in *Philosophical Studies*.

Abstract

In this paper, we refine Lewis's regularity theory of causation by causal models. We show that the resulting theory overcomes the problems that speak decisively against his regularity theory. Further refinements address issues with the transitivity of causation and isomorphic causal scenarios. We conclude that the final theory can compete with the most advanced regularity and counterfactual accounts of causation.
Keywords. Causation, Regularity Theory, Counterfactual Accounts, Causal Models.

1 Introduction

Causation is instantiation of regularities. This is the core idea behind the regularity theory of causation. Lewis (1973) authored a regularity theory just for the purpose of criticising and rejecting it. On this theory, a cause is an indispensable member of any minimal set of actual conditions which jointly entail the effect in the presence of the laws. If so, we say for brevity that the effect is *inferable* from the cause. The theory cannot distinguish genuine causes from effects and preempted would-be causes. These problems speak decisively against it.

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

In this paper, we refine Lewis's regularity theory by embedding it into a framework of causal models. The idea behind our theory is this: causation is *forward-directed inferability along lawful paths*. The forward-directedness overcomes the problems that speak decisively against Lewis's regularity theory. Our idea of lawful paths addresses causal scenarios which suggest that causation is not transitive. It says, roughly, that an effect must be inferable from a genuine cause in the presence of all and only the lawful paths between them. Moreover, we offer an optional condition of deviancy to address the problem of isomorphic causal models: there are pairs of scenarios which are structurally indistinguishable for simple causal model accounts, and yet our causal judgments differ (Hall, 2007). We conclude that our regularity theory can compete with the most advanced regularity and counterfactual accounts of causation.

We proceed as follows. First, we introduce the regularity theory authored by Lewis and the problems it faces. Second, we embed this theory of deterministic token causation into a simple framework of causal models, add the requirement of forward-directedness, and show how the resulting theory overcomes the problems. Third, we further refine the theory by our idea of lawful paths and offer an optional condition of deviancy. Finally, we compare our theory to both other regularity theories and counterfactual accounts.

2 Lewis's Regularity Theory

Hume (1748/1975, Sect. VII) said that causation is the instantiation of regularity. This core idea of the regularity theory has already been refined before Lewis had any chance to criticise it. We learned from Mill (1843/2011) and others that causation requires the instantiation of a specific kind of regularity: laws of nature. Mere accidental regularities do not establish genuine causal relations. Since authors like Hart and Honoré (1959/1985) and Mackie (1965) we allow one indispensable condition to be a cause as long as the totality of conditions is invariably followed by the effect according to at least one law. The regularity theory thus counts as a cause each indispensable member of any minimal set of actual conditions which are jointly sufficient for the effect to occur in the presence of the laws.

Lewis (1973, p. 556) made the regularity theory more precise. He understands sufficiency as entailment: a set of conditions is sufficient for an effect to occur just in case the set entails the effect in the sense of classical logic. On his regularity theory, an event c is thus a cause of another event e if and only if (iff) c belongs to a minimal set of actual conditions that entail the occurrence of e in the presence of the laws. If so, we say e is inferable from c for short.

Here is Lewis's statement of the regularity theory. Let A be the proposition that is true if and only if (iff) the token event a occurs. Furthermore, let $\mathcal{L}$ denote a set of law-like propositions entailed by the true laws and $\mathcal{F}$ a possibly empty set of true propositions of particular fact.

> c is a cause of e iff there is a set $\mathcal{F}$ of true propositions of particular fact and a set $\mathcal{L}$ of true law-like propositions such that all of the following conditions are satisfied:
>
> (1) C and E are true.
>
> (2) $\mathcal{L} \cup \mathcal{F} \models C \to E$.
>
> (3) $\mathcal{L} \cup \mathcal{F} \not\models E$.
>
> (4) $\mathcal{F} \not\models C \to E$.

Let us explain this regularity theory. (1) says that cause and effect are actual. (2) says that a cause entails its effect in the presence of $\mathcal{L} \cup \mathcal{F}$. However, (3) says that $\mathcal{L} \cup \mathcal{F}$ alone does not entail E. Given $\mathcal{L} \cup \mathcal{F}$, C is indispensable for E. In this sense, $\mathcal{L} \cup \mathcal{F} \cup \{C\}$ is a *minimal* set which entails E. (4) says that $\mathcal{F}$ alone does not entail the material implication $C \to E$. This is Lewis's way to implement that the set $\mathcal{L}$ of true law-like propositions is not redundant for the entailment of E.

The set $\mathcal{F}$ contains only propositions of particular fact. The negation $\neg A$ of an actual event a, for example, cannot be in it. It follows from (1)-(4) that the possibly empty set $\mathcal{F}$ alone neither entails C nor E. If it alone were to entail E, (4) would be violated. If it alone were to entail C, either (2) or (3) would be violated. Finally, note that the usage of the material implication is not essential. By the deduction theorem of classical logic, clauses (2) and (4) can be equivalently rephrased as follows:

(2′) $\mathcal{L} \cup \mathcal{F} \cup \{C\} \models E$, and

(4′) $\mathcal{F} \cup \{C\} \not\models E$.

The presented regularity theory faces a problem: it recognizes more causes than there are. It wrongly counts as causes (a) effects of unique causes, (b) joint effects of common causes, and (c) preempted would-be causes.

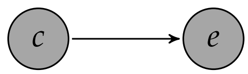

Figure 1

(a) The problem of unique causes. If c is inferable from e, e may nevertheless be an effect of c rather than a cause. Consider the scenario depicted in Figure 1, where c causes e, but e does not cause c, and there are no other causes for e. In this scenario, the law-like propositions $\mathcal{L}$ entail the bi-implication $C \leftrightarrow E$. And so $\mathcal{L}$ and the empty $\mathcal{F}$ entail the implication $E \rightarrow C$ going against the direction of causation. The empty $\mathcal{F}$ neither entails C nor E. Hence, the clauses (1)-(4) are satisfied.

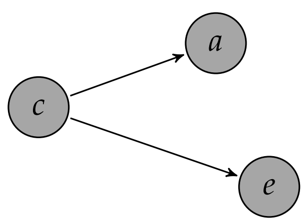

Figure 2

(b) The problem of joint effects. If e is inferable from a, a and e may be joint effects of a common cause c. Consider the scenario depicted in Figure 2, where c causes a and e, but a does not cause e and e does not cause a. Furthermore, a could not have been caused otherwise than by c and c could not have failed to cause e. In this scenario, the law-like propositions $\mathcal{L}$ entail $C \leftrightarrow A$ and $C \rightarrow E$. And so $\mathcal{L}$ and the empty $\mathcal{F}$ entail $A \rightarrow C$ against the direction of causation, and $C \rightarrow E$ in the direction of causation.

By the transitivity of the material implication, we obtain $A \rightarrow E$. Hence, the clauses (1)-(4) are satisfied and a counts as a cause of e.

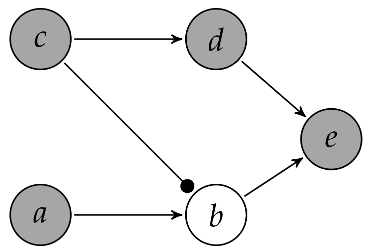

Figure 3

(c) The problem of preemption. If e is inferable from a, a may be a mere would-be cause of e. Consider the scenario depicted in Figure 3, where a did not cause e but would have had the genuine cause c been absent. In this scenario, the law-like propositions entail $A \rightarrow E$. There is some $\mathcal{F}$, which does not contain anything that implies A, E and/or C, such that the clauses (1)-(4) are satisfied. Hence, the mere would-be cause a counts as a cause of e although the genuine cause c preempted the causal efficacy of a.

What went wrong? We propose that an effect must be inferable from a genuine cause in a *causally forward-directed way*. As we have seen in the problem of unique causes and the problem of joint effects, entailments against the direction of causation lead to the recognition of causal relations where there are none. Mere inferability of some event from a putative cause is not enough.

Moreover, one element of regularity theories is the respect for true particular facts. But the entailment of $A \rightarrow E$ in the problem of preemption involves an inference via B, even though b does not occur. This suggests that the regularity theory of Lewis is too liberal as to the choice of the set $\mathcal{F}$ of true propositions of particular fact. There is a minimality constraint on $\mathcal{F}$, but a maximality constraint is lacking which would guarantee that $\neg B \in \mathcal{F}$.

Such a maximality constraint alone, however, does not help Lewis's regularity theory. Even if $\mathcal{F} = \{\neg B\}$, a counts as a cause of e in the preemption scenario. For this to be seen, observe first that the law-like propositions in $\mathcal{L}$ and $\neg B$ entail $\neg A \vee C$. There are only two cases: if $\neg C$, then $\mathcal{L}$ and $\mathcal{F}$ entail $\neg A$, and thus $A \rightarrow E$. If C, then $\mathcal{L}$ and $\mathcal{F}$ entail E, and thus $A \rightarrow E$. Condition (2) is again satisfied. But observe that this reasoning is artificial. Intuitively, a is not a cause of e because A does not entail E in a forward-directed way via B.

In the next section, we embed the regularity theory authored by Lewis into a framework of causal models. This allows us to add both: a maximality constraint on the minimal set of actual conditions which jointly entail the effect, and a requirement of forward-directedness on the inferability of the effect.

3 Refining Lewis's Regularity Theory

We refine Lewis's regularity theory by embedding it into a framework of causal models. For this purpose, we introduce first a simple framework of causal models and define an entailment relation in causal models. We then analyse causation in the spirit of Lewis's regularity theory while taking the lessons from the last section into account: causation requires a condition of forward-directedness and a maximality constraint. Finally, we revisit the three troublesome causal scenarios.

3.1 A Framework of Causal Models

Causal models represent causal scenarios. In a causal scenario like preemption, certain events occur, others do not, and we have a certain law-like structure that tells us how event types depend on other event types. We define a causal model $\langle \mathcal{L}, \mathcal{F} \rangle$ by two components: a set $\mathcal{L}$ of law-like propositions and a set $\mathcal{F}$ of true propositions of particular fact. $A \in \mathcal{F}$ means that some token event a of type A occurs. $\neg A \in \mathcal{F}$ means that no token event of type A occurs. In other words, $\neg A$ denotes the absence of any event of type A, or simply the absence of A.

A law-like proposition has the form

$$A = \phi,$$

where A is a propositional variable, ϕ a propositional formula in disjunctive normal form, and no variable appears vacuously. So 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 A or $\neg A$. A law-like proposition expresses the true regularity that A iff ϕ. We say a propositional variable appears in $A = \phi$ vacuously iff the variable never affects the truth values of A and ϕ. In the law-like proposition $A = C \vee (D \wedge \neg D)$, for example, the variable D appears vacuously.

In our framework, law-like propositions are directed bi-implications. They have a variable A standing for a type effect on the left-hand side and a Boolean combination of variables standing for type causes on the right-hand side. We take the direction of law-like propositions as given. As a consequence, our theory is not reductive. We discuss the prospects of a reductive regularity theory in Section 6.1.

Assuming the direction of law-like propositions, the preemption scenario can be represented by a causal model $\langle \mathcal{L}, \mathcal{F} \rangle$, where $\mathcal{L} = \{D = C, B = A \wedge \neg C, E = D \vee B\}$ and $\mathcal{F} = \{C, A, D, \neg B, E\}$. For readability, we represent causal models in two-layered boxes. The upper layer shows the set $\mathcal{L}$ of law-like propositions. The lower layer shows the set $\mathcal{F}$ of propositions of particular fact. For the preemption scenario, we obtain:

$D = C$ $B = A \wedge \neg C$ $E = D \vee B$
$C, A, D, \neg B, E$

Let us define a causal model semantics in terms of the semantics of propositional logic. We say a classical valuation satisfies a law-like proposition $A = \phi$ iff both sides have the same truth value on this valuation. This allows us to define the satisfaction relation in the standard way. Where Γ is a set of propositional formulas and law-like propositions, $\Gamma \models \psi$ iff the propositional formula or law-like proposition ψ is satisfied by any classical valuation that satisfies all members of Γ. We define the entailment relation in causal models as follows:

$$\langle \mathcal{L}, \mathcal{F} \rangle \models \psi \quad \text{iff} \quad \mathcal{L} \cup \mathcal{F} \models \psi.$$

Finally, we say that a set Γ of propositional formulas and law-like propositions satisfies another such set Δ iff $\Gamma \models \psi$ for any ψ in Δ.

A central idea of our theory is that an effect is inferable from its cause in a forward-directed way. A law-like proposition $A = \phi$ has the truth conditions of the bi-implication $A \leftrightarrow \phi$ and so is symmetric: it allows for forward-directed inferences from ϕ to A and backward-directed inferences from A to ϕ. We introduce the notion of a *setting* to isolate the forward-directed causal consequences of some event a of type A for a causal model $\langle \mathcal{L}, \mathcal{F} \rangle$. Roughly speaking, a setting removes a law-like proposition $A = \phi$ from $\mathcal{L}$ and replaces it by a true proposition, either A or $\neg A$. Thereby backward-directed inferences from A or $\neg A$ are excluded.

Settings establish an asymmetry based on the direction of law-like propositions. Consider, for example, a causal model which includes the law-like proposition $E = C$. Setting C determines E in a forward-directed way. However, setting E does not determine C, it removes the law-like proposition and replaces it by E. The considered law-like proposition has the same truth conditions as $C = E$. But had the latter instead of the former been in the causal model, setting C would have removed this law-like proposition and setting E would have determined C in a forward-directed way. The difference between $E = C$ and $C = E$ matters for what is and isn't inferable in a forward-directed way. In general, the direction of the law-like propositions matters for the direction of causation.

Suppose we want to determine the forward-directed causal consequences of the occurring token event a of type A for a causal model $\langle \mathcal{L}, \mathcal{F}' \rangle$. The setting of A in this causal model results in a causal model $\langle \mathcal{L}_A, \mathcal{F}' \cup \{A\} \rangle$. If $A = \phi$ is a member of $\mathcal{L}$, $\mathcal{L}_A$ is obtained from $\mathcal{L}$ by removing this law-like proposition. Otherwise $\mathcal{L}_A = \mathcal{L}$. We call $\langle \mathcal{L}_A, \mathcal{F}' \cup \{A\} \rangle$ the causal submodel of $\langle \mathcal{L}, \mathcal{F}' \rangle$ after the setting of A. By removing the law-like proposition of A from $\mathcal{L}$, backward-directed inferences from A or $\neg A$ are excluded in the causal submodel. The asymmetry of causation may so be established by a setting and the direction of the law-like propositions.

In general, we denote possibly complex settings by an operator $[\cdot]$ that takes a causal model $\langle \mathcal{L}, \mathcal{F}' \rangle$ and a set $\mathcal{S}$, where both $\mathcal{F}'$ and $\mathcal{S}$ are subsets of the true propositions $\mathcal{F}$ of particular fact, and returns a causal model: the submodel of $\langle \mathcal{L}, \mathcal{F}' \rangle$ after the setting of $\mathcal{S}$. The setting by a set of true propositions of particular fact is defined as follows:

$$\langle \mathcal{L}, \mathcal{F}' \rangle[\mathcal{S}] = \langle \mathcal{L}_{\mathcal{S}}, \mathcal{F}' \cup \mathcal{S} \rangle$$

where

$$\mathcal{L}_{\mathcal{S}} = \{(A = \phi) \in \mathcal{L} \mid A \notin \mathcal{S} \text{ and } \neg A \notin \mathcal{S}\}.$$

$\mathcal{L}_{\mathcal{S}}$ is the subset of $\mathcal{L}$ that contains each law-like proposition $A = \phi$ whose variable A does not appear in $\mathcal{S}$. After setting $\mathcal{S}$ in the causal model $\langle \mathcal{L}, \mathcal{F}' \rangle$, the set $\mathcal{S}$ becomes part of the propositions of particular fact of the resulting submodel. Note that the resulting submodel is again a causal model consisting of a set of law-like propositions and a set of propositions of particular fact.

Settings will always only set true propositions of particular fact. No propositions contrary to the true facts are ever set, unlike the interventions employed by Halpern and Pearl (2005) for example. As a consequence, the submodels resulting from settings are not inconsistent provided the original causal models were not.

Our directed law-like propositions resemble symmetric structural equations. For some authors, the structural equations themselves are asymmetric and thereby exclude inferences against the direction of causation (Hitchcock, 2001). For others the asymmetry comes in only through the interventions defined for structural equations (Pearl, 2009). For us the asymmetry comes in only through the settings defined for directed law-like propositions.

3.2 A Refined Regularity Theory

We are now in a position to refine the regularity theory of causation.

Definition 1. Let $\langle \mathcal{L}, \mathcal{F} \rangle$ be a causal model such that $\mathcal{F}$ satisfies $\mathcal{L}$. c is a cause of e relative to $\langle \mathcal{L}, \mathcal{F} \rangle$ iff there is a possibly empty set $\mathcal{F}' \subseteq \mathcal{F}$ such that all of the following conditions are satisfied:

(i) $\langle \mathcal{L}, \mathcal{F} \rangle \models C \wedge E$.

(ii) $\langle \mathcal{L}, \emptyset \rangle [\mathcal{F}'][\{C\}] \models E$.

(iii) $\langle \mathcal{L}, \mathcal{F}' \rangle \not\models E$ and there is no $\mathcal{F}''$ so that $\mathcal{F}' \subset \mathcal{F}'' \subseteq \mathcal{F}$ and $\langle \mathcal{L}, \mathcal{F}'' \rangle \not\models E$.

(i) says that cause and effect are actual. (ii) says that, in the presence of the law-like propositions $\mathcal{L}$, a cause together with some propositions $\mathcal{F}'$ of particular fact entails its effect in a forward-directed way. However, (iii) says that the propositions $\mathcal{F}'$ of particular fact and the law-like propositions $\mathcal{L}$ alone do not entail E; and it requires that $\mathcal{F}'$ is maximal: any strict superset of $\mathcal{F}'$ would entail E in the presence of the law-like propositions.

Our preliminary regularity theory resembles the regularity theory authored by Lewis. $\mathcal{F}'$ is *some* set of true particulars such that the effect proposition E is forward-directedly entailed by it together with a genuine cause proposition C in the presence of the law-like propositions in $\mathcal{L}$; and yet $\mathcal{F}'$ and $\mathcal{L}$ alone do not entail E. C is indispensable for the forward-directed entailment.

However, our preliminary regularity theory is stronger than Lewis's. (ii), as compared to (2), is strengthened by the requirement of forward-directedness. (iii), as compared to (3), is strengthened by a maximality condition that implements a respect for the true particular facts. A genuine cause proposition C is thus an indispensable member of a minimal set of actual conditions that entail E in a forward-directed way, while it contains as many as possible of the actual facts. The two strengthenings make an equivalent to Lewis's condition (4) superfluous.

On our preliminary theory, a cause is each member of any *maximised* minimal set of actual conditions which, in the presence of the law-like propositions, entail the effect in a *forward-directed* way. Causation so understood is lawful inferability in a forward-directed way that respects the particular facts. It is time to revisit the troublesome causal scenarios.

3.3 Causal Scenarios Revisited

The refined regularity theory gives the correct verdicts for the three troublesome scenarios we have considered so far. Consider the causal model $\langle \mathcal{L}, \mathcal{F} \rangle$ for the *problem of unique causes*:

$E = C$
C, E

Here, c is a cause of e. C and E are true in the causal model, and (ii) and (iii) are satisfied for $\mathcal{F}' = \emptyset$.

By contrast, e is not a cause of c. There is no $\mathcal{F}'$ that satisfies (ii) and (iii). (ii) demands that $\langle \mathcal{L}, \emptyset \rangle [\mathcal{F}'][\{E\}]$ entails C. The setting of E removes the law-like proposition $E = C$ from $\mathcal{L}$. (ii) is then only satisfied if $\mathcal{F}'$ contains C. But then $\langle \mathcal{L}, \mathcal{F}' \rangle \models E$ which violates (iii). Indeed, c is not inferable from e in a forward-directed way.

Consider the causal model $\langle \mathcal{L}, \mathcal{F} \rangle$ for the *problem of joint effects*:

$A = C$ $E = C$
C, A, E

Here, c is a cause of e. C and E are true in the causal model, and (ii) and (iii) are satisfied for $\mathcal{F}' = \emptyset$. Similarly, c is a cause of a.

By contrast, a is not a cause of e. There is no $\mathcal{F}'$ that satisfies (ii) and (iii). (ii) demands that $\langle \mathcal{L}, \emptyset \rangle [\mathcal{F}'][\{A\}]$ entails E. The setting of A removes the law-like proposition $A = C$ from $\mathcal{L}$. (ii) is then only satisfied if $\mathcal{F}'$ contains C or E. In both cases $\langle \mathcal{L}, \mathcal{F}' \rangle \models E$, which violates (iii). Indeed, e is not inferable from a in a forward-directed way. Similarly, e is not a cause of a.

The problems of unique causes and joint effects illustrate how settings establish the asymmetry of causation based on the direction of law-like propositions. In the presence of the law-like proposition $A = \phi$, a setting of some proposition in ϕ may determine whether A or $\neg A$, but a setting of A does not determine any truth value of any proposition appearing in

ϕ. In the presence of settings, any $A = \phi$ says that ϕ determines whether A or $\neg A$ in a forward-directed way and not the other way around. We use this feature of settings to identify the direction of causation. A correct identification is necessary to solve the problems of unique causes and joint effects. This means: our theory can solve these problems only if we have identified the true law-like propositions and their direction.

Consider the causal model $\langle \mathcal{L}, \mathcal{F} \rangle$ for the *problem of preemption*:

$D = C$
$B = A \land \neg C$
$E = D \lor B$
$C, A, D, \neg B, E$

Here, c is a cause of e. C and E are true in the causal model, and (ii) and (iii) are satisfied for $\mathcal{F}' = \{\neg B\}$.

By contrast, a is not a cause of e. There is no $\mathcal{F}'$ that satisfies (ii) and (iii). (iii) demands that $\langle \mathcal{L}, \mathcal{F}' \rangle \not\models E$ and every strict superset of $\mathcal{F}'$ that is a non-strict subset of $\mathcal{F}$ would entail E. So $\mathcal{F}'$ must be the set $\{\neg B\}$. (ii) then demands that $\langle \mathcal{L}, \emptyset \rangle [\{\neg B\}][\{A\}]$ entails E. But this is not the case.

We have shown this: once we have the true law-like propositions and their direction, our preliminary regularity theory overcomes the three problems that speak decisively against the regularity theory authored by Lewis. We discuss the extent to which the direction of law-like propositions may be obtained in sections 6.1.1 and 6.1.2. For now, we address further troubles beginning with causal scenarios that threaten the transitivity of causation.

4 Transitivity

Our preliminary regularity theory is challenged by causal scenarios which suggest that causation is not transitive. The transitivity of causation means this: whenever a token event c is a cause of another a, and a is a cause of a third event e, then c is a cause e. It seems often plausible to judge c a cause of e if you judge c a cause of a and a a cause of e. However, several scenarios have been put forth which suggest that our causal judgments are not transitive (McDermott, 1995; Lewis, 2000; Paul, 2000).

4.1 The Boulder Scenario

One of the examples against transitivity may be found in Hitchcock (2001, p. 276). A boulder is dislodged and rolls toward a hiker. The hiker sees the boulder coming and ducks, so that she does not get hit by the boulder. If the hiker had not ducked, however, the boulder would have hit her.

The boulder scenario seems to show that there are cases where causation is not transitive: the dislodged boulder causes the ducking of the hiker, which in turn causes the hiker to remain untouched by the boulder. But the dislodging of the boulder does not cause the hiker to remain unscathed. Unlike other accounts, our theory does not rely on transitivity to handle certain causal scenarios. We are thus free to deny that causation is invariably transitive.

The formal representation of informal stories like the boulder example is somewhat controversial. We think Paul and Hall (2013, pp. 223-6) argue successfully against the causal model for the boulder scenario employed and argued for by Hitchcock (2001, pp. 295-8). Following Gallow (2021, p. 53), we represent the structure of the boulder scenario as follows.

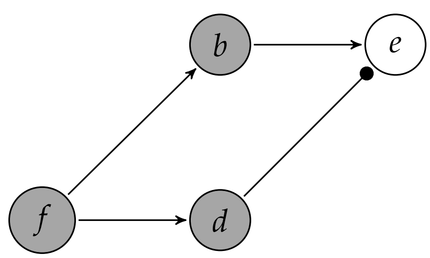

Figure 4

Hall (2007, p. 36) calls this structure a *short circuit*: the boulder's dislodgement f threatens to hit the hiker by a rolling boulder b, and at the same time provokes an action—the ducking d—that prevents this threat from being effective: no token event e of type E occurs.

The token event f should not count as a cause of the absence of any event of type E, because f creates *and* cancels the threat to bring about some token event of type E (Paul and Hall, 2013, p. 216). Our preliminary regularity theory says so. However, it does also not count the ducking of the hiker as a cause of her remaining unscathed: d does not count as a cause of $\neg E$, or 'e's absence'. To see this, consider the causal model of the boulder scenario:

$B = F$
$D = F$
$E = B \wedge \neg D$
$F, B, D, \neg E$

Note that the set $\mathcal{L}$ of law-like propositions alone entails $\neg E$—even if $\mathcal{F}' = \emptyset$. $\neg E$ is entailed whether F is true or not, and in violation of condition (iii). The verdict that the ducking of the hiker is not a cause of the hiker's remaining untouched is clearly false. And so our preliminary theory succumbs to a severe problem.

In response, we propose to further refine our theory by an idea of *lawful paths between cause and effect*. The basic idea is that an effect must be inferable from a genuine cause in the presence of all and only the lawful paths between them. The lawful paths between cause and effect are just a set of law-like propositions that connect cause and effect in a forward-directed way. One may roughly think of lawful paths as chains of lawful propositions running from a candidate cause to its putative effect. The idea of lawful paths adds this to our regularity theory: when testing whether a token event c of type C is a cause of a token event e of type E, the lawful paths from C to E need to remain intact, but only those paths.

We implement the idea of lawful paths between cause and effect in two steps. First, we allow to remove law-like propositions when testing for causation. Second, we constrain the removal of law-like propositions by a condition which ensures that the lawful paths starting from a candidate cause remain in the law-like propositions.

Our implementation requires some terminology. We say $A = \phi$ is the law-like proposition of A. We also say A is a *child* variable of the *parent* variables appearing in ϕ. Let B be one of the parent variables appearing in ϕ. The variable A is then a child variable of B, and so a first descendant of B. The child variables of A are the child variables of one of B's child variables, and so are among B's second descendants. In general, the descendants of some variable B are the variables in the transitive closure of the child relation starting from B. Similarly, the ancestors of some variable B are the variables in the transitive closure of the inverse child relation—the parent relation—starting from B.

Where B is a propositional variable, let D be a proposition of the form B or $\neg B$. We say that the descendants of the proposition D are all the variables (of the causal model under consideration) which are descendants of the variable B. We are now in a position to state our regularity theory of causation.

Definition 2. Let $\langle \mathcal{L}, \mathcal{F} \rangle$ be a causal model such that $\mathcal{F}$ satisfies $\mathcal{L}$. c is a cause of e relative to $\langle \mathcal{L}, \mathcal{F} \rangle$ iff there are possibly empty sets $\mathcal{F}' \subseteq \mathcal{F}$ and $\mathcal{L}' \subseteq \mathcal{L}$ such that all of the following conditions are satisfied:

(i) $\langle \mathcal{L}, \mathcal{F} \rangle \models C \wedge E$.

(ii) $\langle \mathcal{L}', \emptyset \rangle [\mathcal{F}'][\{C\}] \models E$.

(iii) $\langle \mathcal{L}', \mathcal{F}' \rangle \not\models E$ and there is no $\mathcal{F}''$ so that $\mathcal{F}' \subset \mathcal{F}'' \subseteq \mathcal{F}$ and $\langle \mathcal{L}', \mathcal{F}'' \rangle \not\models E$.

(iv) For all descendants A of C, the law-like proposition of A is in $\mathcal{L}'$.

The removal of law-like propositions is implemented as follows. We allow a subset $\mathcal{L}'$ of $\mathcal{L}$ to figure in the condition (ii) of forward-directed inferability and condition (iii) of C's indispensability for the inferability in a maximal context of actual facts. The removal of law-like propositions is constrained by condition (iv): the causal paths starting from a candidate cause C must remain in $\mathcal{L}'$. Causation so understood is *forward-directed inferability along the lawful paths between cause and effect*.

Let us reconsider the boulder example. The dislodgement of the boulder f still does not count as a cause of the hiker's remaining unscathed $\neg E$. To see this, observe that all variables of the causal model are descendants of F. Condition (iv) thus ensures that all law-like propositions of $\mathcal{L}$ must remain in $\mathcal{L}'$. As a consequence, there is no $\mathcal{L}'$ and $\mathcal{F}'$ so that condition (iii) is satisfied. Even for $\mathcal{F}' = \emptyset$, the law-like propositions entail $\neg E$.

The hiker's ducking d, by contrast, counts as a cause of the hiker's remaining unscathed $\neg E$. To see this, observe that F is not a descendant of D. Hence, the law-like proposition $D = F$ can be removed from $\mathcal{L}$. Take $\mathcal{L}' = \{B = F, E = B \wedge \neg D\}$ and $\mathcal{F}' = \{F, B\}$. $\langle \mathcal{L}', \mathcal{F}' \rangle$ does not entail $\neg E$, and $\mathcal{F}'$ is maximal: any strict superset of $\mathcal{F}'$ would entail $\neg E$ in the presence of the law-like propositions in $\mathcal{L}'$. But, of course, $\langle \mathcal{L}', \emptyset \rangle [\mathcal{F}'][\{D\}] \models \neg E$.

We have further refined our regularity theory by an idea of lawful paths between cause and effect. The refined regularity theory says that causation is not invariably transitive. In the boulder scenario, the dislodged boulder comes out as a cause of the hiker's ducking, which in turn comes out as a cause of the hiker's remaining unscathed. And yet, the dislodged boulder does not count as a cause of the hiker's remaining unscathed. We turn now to further examples which suggest that causation is not transitive.

4.2 Simple Switch

Switching scenarios are paradigmatic for causal scenarios where our causal judgments are not transitive. In switching scenarios, some occurring event of type F helps determine the causal path by which another event is brought about. Crucially, the other event would also occur via an alternative causal path if no event of type F had occurred.

To make it more concrete, consider a story provided by Hall (2000, p. 205). Flipper is standing by a switch in the railroad tracks. A train approaches in the distance. She flips the switch, so that the train travels down the right track, instead of the left. Since the tracks reconverge up ahead, the train arrives at its destination all the same. The commonsense judgment is that flipping the switch is not a cause of the train's arrival—even though flipping the switch is a cause of the train's travelling on the right track, and the train's travelling on the right track is a cause of the train's arrival (Paul, 2000; Yablo, 2002; Sartorio, 2005, 2006; Schaffer, 2005; Hall, 2007; Hitchcock, 2009; Paul and Hall, 2013; Baumgartner, 2013; Halpern, 2016; Beckers and Vennekens, 2018; Andreas and Günther, 2021b; Gallow, 2021).

The structure of this simple switch scenario can be visualized by the following diagram (Beckers and Vennekens, 2018, p. 846).

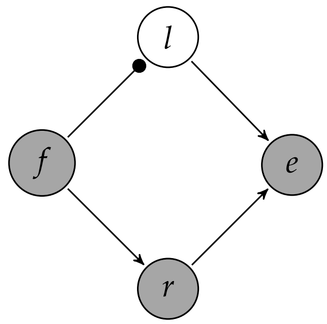

Figure 5

The flipping of the switch f causes the train to travel on the right track r and prevents the train from travelling on the left track l. And the travelling on the right track r causes the train to arrive at its destination e. However, the flipping of the switch f arguably is not a cause of the train's arrival e. Here is the causal model for the switch scenario:

$L = \neg F$ $R = F$ $E = L \vee R$
$F, \neg L, R, E$

Relative to this causal model, f is not a cause of e. For this to be seen, observe first that all variables are descendants of F. Condition (iv) thus prohibits to remove any law-like proposition from $\mathcal{L}$. But then there is no $\mathcal{F}'$ so that condition (iii) is satisfied. Even for $\mathcal{F}' = \emptyset$, the law-like propositions entail E.

By contrast, f is a cause of r. Take $\mathcal{L}' = \mathcal{L}$ and $\mathcal{F}' = \{E\}$. Condition (iii) is then satisfied: $\langle \mathcal{L}', \mathcal{F}' \rangle \not\models R$, and $\mathcal{F}'$ is maximal—any strict superset of $\mathcal{F}'$ would entail R in the presence of the law-like propositions $\mathcal{L}'$. The other conditions are trivially satisfied.

Likewise, r is a cause of e. Take $\mathcal{L}' = \{L = \neg F, E = L \vee R\}$ and $\mathcal{F}' = \{F, \neg L\}$. Condition (iii) is then satisfied because $\langle \mathcal{L}', \mathcal{F}' \rangle \not\models E$; and $\mathcal{F}'$ is maximal—any strict superset of $\mathcal{F}'$ would entail E in the presence of the law-like propositions $\mathcal{L}'$. The other conditions are trivially satisfied.

In sum, our regularity theory delivers the desired verdicts in both the boulder and simple switch scenarios. The representation of switch scenarios is somewhat controversial. The diagram in Figure 5 should not be read as a neuron diagram, but rather as a causal model. We will discuss further switch scenarios in Section 6.3. For now, we turn towards the problem of isomorphic causal models.

5 Isomorphic Causal Models

The problem of isomorphic causal models is that there are pairs of scenarios which are structurally indistinguishable for simple causal model accounts, and yet our causal judgments differ (Hall, 2007, p. 44). We call a causal model account simple if it only factors in structural equations—or our law-like propositions—together with values of variables—or our propositions of particular fact.

5.1 Bogus Prevention

Let us illustrate an instance of the problem. Consider a scenario of overdetermination, where an effect is overdetermined by more than one event. An example runs as follows: a prisoner is shot e by two soldiers c and a at the same time, and each of the bullets is fatal without any temporal precedence. Each of the shots is a cause of the death of the prisoner. The structure of this scenario, where the effect e is overdetermined by the two causes c and a, can be represented as follows.

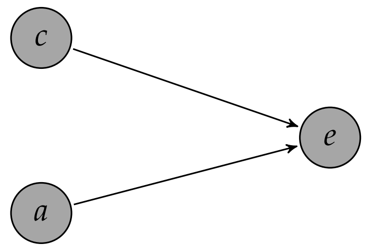

Figure 6

Here is the causal model of the overdetermination scenario.

$E = C \vee A$
C, A, E

Our regularity theory says c is a cause of e, and so is a. So far so good. The problem is now that this causal model can be transformed into a structurally indistinguishable or isomorphic one, for which our causal judgment differs. The transformation, first, negates both sides of the law-like proposition. Then it substitutes C by F, A by $\neg D$, and E by $\neg E$. The result is the isomorphic causal model:

$E = \neg F \wedge D$
$F, \neg D, \neg E$

And indeed, $\neg E$ is 'overdetermined' by F and $\neg D$. The isomorphic causal model can be visualized as follows.

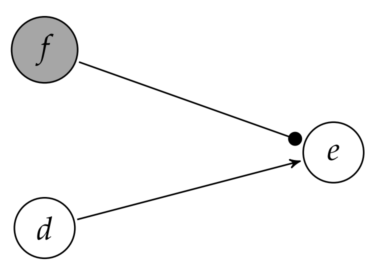

Figure 7

Here is a story for this structure. There is an assassin, a potential target, and her bodyguard. The assassin refrains from poisoning target's coffee $\neg D$, and her bodyguard puts antidote in her coffee f. Target survives $\neg E$, of course. Since target's coffee is not poisoned in the first place, there is no danger at all that she dies. The prevention by bodyguard's antidote is *bogus*. And so bodyguard's putting the antidote in her coffee is arguably no cause of her survival (Hiddleston, 2005; Hitchcock, 2007).

Recall that simple causal model accounts cannot detect any difference between the administration of antidote f in the scenario of bogus prevention and the shot c of one of the soldiers in the scenario of overdetermination. However, our causal judgments differ. We judge c to be a cause of e in the overdetermination scenario, while we do not judge f to be a cause of $\neg E$ in the bogus prevention scenario.

Simple causal model accounts of causation—for instance the accounts of Hitchcock (2001) and Halpern and Pearl (2005)—cannot distinguish between f and c in the isomorphic causal models: c counts as a cause iff f does. This means simple causal model accounts must incorrectly classify f as a cause in the bogus prevention scenario if they correctly classify c as a cause in the overdetermination scenario. This is a problem indeed.

Our theory of causation is a simple causal model account, and so is likewise susceptible to the problem of isomorphic causal models. A solution favored by many authors relies on default or normality considerations (Hitchcock, 2007; Hall, 2007; Halpern and Hitchcock, 2015; Halpern, 2015). The underlying idea is that the status of genuine causes depends on being deviant from what is normal (Beebee, 2004; McGrath, 2005). On this view, genuine effects are brought about by causes that are more deviant from normality than its non-actual alternatives. Gallow (2021) goes even further by elevating the transmission of deviancy to the mark of causation: an event is a cause in virtue of transmitting its deviancy to its effect.

We can also resolve the problem of isomorphic causal models by a condition of deviancy. But what is deviancy? For now, we follow Gallow (2021, p. 54) in saying that *prima facie* an occurring event is more deviant than its absence. But this is just a first approximation. The question of what constitutes deviancy is more intricate, as we will see below.

We amend now our theory by an optional condition of deviancy. The condition is motivated by the idea that any cause of an effect must be deviant. We implement the idea as follows: any candidate cause C' of an effect E, which is neither a descendant nor an ancestor of the candidate cause C under consideration, must be deviant. Note that C is neither a descendant nor an ancestor of itself. Any one C' in $\mathcal{F} \setminus \mathcal{F}'$ is a candidate cause of E because it entails the effect together with the propositions in $\mathcal{F}'$ in the presence of the laws $\mathcal{L}'$. Otherwise C' would remain in $\mathcal{F}'$ in virtue of its maximality. To be precise, we offer the option to add the following condition to conditions (i)-(iv):

(v) for any proposition C' in $\mathcal{F} \setminus \mathcal{F}'$ whose variable is neither a descendant nor an ancestor of C, C' is more deviant than $\neg C'$.

The deviancy condition (v) says that, for c to be a cause of e, the proposition C and each proposition C', which may form with $\mathcal{F}'$ some maximised minimal set for E and whose variable is neither a descendant nor an ancestor of C, must be more deviant than its respective negation. It follows that the propositions along the lawful paths from each C' to E may be non-deviant so long as any cause C is deviant. On our so-amended regularity theory, causation is understood as *forward-directed inferability along lawful paths from deviant events and absences*.

The above scenario of bogus prevention illustrates the underlying rationale of the deviancy condition (v). The absence of poison $\neg D$ and the presence of antidote f in the coffee are no causes because it is normal that coffees are not poisoned: $\neg D$ is not deviant. The normality of non-occurrence of any event of type D entails that 'd's absence' and f's occurrence are no causes of target's survival. The non-deviancy of $\neg D$ robs both itself and f of its causal status. Our regularity theory amended by our condition of deviancy solves both overdetermination and bogus prevention.

5.2 Omissions

Omissions pose another problem for many theories of causation. In a scenario of omission, an event fails to occur and so another event occurs.

However, had the event occurred, it would have prevented the other event from occurring. The basic structure of omissions can be represented as follows.

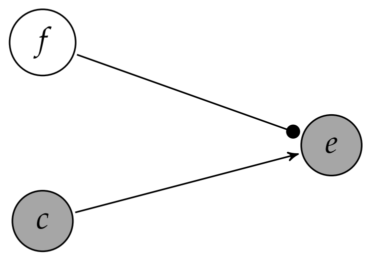

Figure 8

An event c occurs and brings about an event e. No event f of type F occurs. However, had an event f of type F occurred, it would have prevented e from occurring. Here is the causal model for the scenario of omission.

$E = \neg F \wedge C$
$\neg F, C, E$

Relative to this causal model, 'f's absence' is not a cause of e, given the above convention about deviancy. For this to be seen, observe that conditions (ii) and (iii) are only satisfied for $\mathcal{F}' = \{C\}$. But then condition (v) is violated: $\neg F$ is in $\mathcal{F} \setminus \mathcal{F}'$ and its variable F is neither a descendant nor an ancestor of itself, and yet $\neg F$ is less deviant than F. (By contrast, c is a cause of e. Take $\mathcal{F}' = \{\neg F\}$. Conditions (ii) and (iii) are then satisfied. And since $\neg F$ is not in $\mathcal{F} \setminus \mathcal{F}'$, condition (v) is trivially met.)

Indeed, many omissions are no causes. Putin's failure to water my plant, for example, did not cause it to dry up and die. However, some omissions intuitively *do* count as causes. My neighbour promised me to water my plant, but she didn't and it died. Here my neighbour's failure to water my plant should count as a cause of its death (McGrath, 2005). Our theory can capture this phenomenon if we refine our notion of deviancy.

We have said that *prima facie* an occurring event is more deviant than its absence. We say now in addition that the absence $\neg A$ of any event of type A

is more deviant than an event of type A if $\neg A$ violates a norm that is active in the scenario under consideration (Beebee, 2004; Andreas et al., 2022). My neighbour's omission to water my plant is an absence that violates the active norm of promise-keeping. My neighbour deviated from this norm and so her omission is more deviant than its negation. Our amended theory says then that my neighbour's failure to water my plant is a cause of the plant's death. Putin, by contrast, did not promise to water my plant. His omission is thus less deviant than his watering my plant, and so does not count as a cause of my plant's death. Or so says our amended theory.

We have illustrated how the condition (v) of deviancy can help to overcome the problem of isomorphic causal models and how it can account for simple scenarios of omission. As to the latter, deviant omissions are genuine causes, non-deviant ones are not. Our regularity theory amended by the condition of deviancy says that a genuine cause is deviant and allows to infer its effect in a forward-directed way along lawful paths. This is a powerful theory if we allow deviancy to play a role in the concept of causation.

6 Comparisons

How does our regularity theory compare to other accounts of causation? In this section, we will locate our theory among other regularity accounts and briefly compare it to counterfactual accounts. We will argue that our theory is compatible with the tradition of 'typical' regularity theories. We will explain Baumgartner's attempt to establish the direction of causation in Section 6.1.1. We do not endorse his attempt for reasons laid out in Section 6.1.2.

We then turn to Wright's (2011) non-reductive regularity account that imposes transitivity on causation. As a consequence, his NESS account faces troubles in scenarios which suggest that causation is not transitive. Several authors have attempted to formalize Wright's NESS account using causal models. We will argue that they either miss their target, or else inherit the problems of Wright's original account, or both.

Finally, we will contrast our regularity theory to counterfactual accounts and discuss another switch scenario due to Halpern and Hitchcock (2010). Our theory can only solve this switch scenario when amended by the optional condition of deviancy.

6.1 'Typical' Regularity Theories

Lewis (1973, p. 556) calls the regularity theory he authored and rejected "typical". And indeed, his proposal reflects the development of the regularity approach to causation until then. The core idea of regularity theories of causation is that causes are regularly followed by their effects. Hume (1748/1975) adds to the instantiation of regularity that a cause is spatiotemporally contiguous to its effect and precedes its effect in time. At least on one reading of Hume, there is nothing more to causation, and so causation is reduced to non-causal entities.

The regularity approach in Hume's tradition aims to be reductive. It is characterised by taking a stance against metaphysically thick conceptions of causation (Dowe, 2000; Psillos, 2002; Andreas and Günther, 2021). The causal relation does, in particular, not involve a necessary connection, a productive relation, unobservable causal powers, or the like—not even to ground the regularities. A regularity is only a stable pattern of events and absences. Cause and effect simply instantiate such a pattern.

Mill (1843/2011) observes that causation requires laws of nature: the most general regularities which subsume all the other true regularities. For Mill (1843/2011, Book I, Ch. V), a cause is a "sum total" of actual conditions which are jointly sufficient for the effect in the presence of the laws of nature. An effect may have many sum totals or sets of conditions that are sufficient for it. Hart and Honoré (1959/1985, p. 112) and Mackie (1965, p. 246) emphasise that each sum total must be minimally sufficient for its effect: without any one of its members, a sum total is not sufficient for its effect. In brief, each member of a sum total is necessary for its sufficiency. And since Hart and Honoré and Mackie, the regularity theory counts each necessary condition of any actual or instantiated sum total a cause.

Mackie (1965, 1974) spells out his theory in terms of complex regularities.

A complex regularity for an effect is a disjunction of conjunctions in disjunctive normal form which is necessary and sufficient for said effect. Here is a toy example of such a complex regularity:

$$(C_1 \wedge C_2) \vee D_1 \leftrightarrow E. \tag{1}$$

The sum total $C_1 \wedge C_2$ is minimally sufficient for the effect E, and so is the sum total D_1. C_1 on its own is insufficient to bring about E. But it is part of the sum total $C_1 \wedge C_2$ which is sufficient but unnecessary for E. Taken together, C_1 is an *i*nsufficient but *n*on-redundant part of an *u*nnecessary but *s*ufficient condition for E. In brief, C_1 is an INUS condition for E.

On Mackie's theory, a token event c is a cause of another e iff C is at least an INUS condition of E and belongs to an instantiated sum total sufficient for E. "At least" because C may also be a necessary, or a sufficient, or even a necessary and sufficient condition for E. This theory says, roughly, that a cause is at least a non-redundant or indispensable member of a minimal set of actual conditions which are jointly sufficient for the effect to occur in the presence of the complex regularities. Lewis (1973) represented this 'typical' regularity theory of his time using the entailment relation of classical logic.

Indeed, like Lewis's statement of the regularity theory, Mackie's succumbs to the problems of unique causes and joint effects. It is controversial whether Mackie's theory solves the problem of preemption. Strevens (2007) argues against Mackie (1974, p. 44-7) that it does. Recall Figure 3. Everyone agrees that c is a cause of e. For c belongs to a set of actual conditions which are jointly sufficient for the effect e to occur, and removing c from that set makes it insufficient. The controversy is whether the event a falsely counts as a cause of e. Strevens says no. Even though A belongs to a minimal sum total $A \wedge \neg C$ sufficient for E, not all conditions of this sum total are actual: c occurs. And he thinks this generalizes to all cases of preemption when sufficiency is replaced by *causal sufficiency*—a notion which we will discuss below.

The underlying problem for Mackie's theory is that it does not give us the direction of causation. The complex regularities are material bi-implications which seem to blur the asymmetry between cause and effect—at least in the problems of unique causes and joint effects.

6.1.1 Non-Redundant Regularities

Baumgartner (2013) developed Mackie's theory further. He observes that complex regularities like (1) show a certain directedness: an instantiation of a sum total, here $C_1 \wedge C_2$ or D_1, is sufficient for E, while an instantiation of E is generally not sufficient to determine which sum total is instantiated. Baumgartner uses this directedness to establish the direction of causation under his *assumption of multiple type causes*: each type effect has at least two type causes.

Here is how Baumgartner aims to establish the direction of causation in a nutshell. The complex regularities must be constrained: they must be rigorously minimized. The left-hand side of each complex regularity must be necessary for its effect in a minimal way. We illustrate this requirement by considering the following joint effects structure: the joint type effects A and B have a common type cause C and each type effect has an alternative type cause, D and E, respectively.

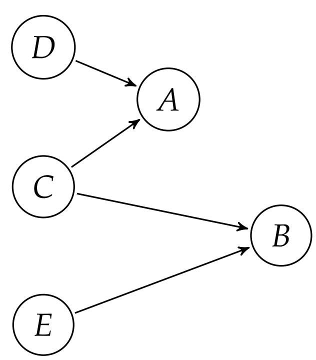

Figure 9

No effect occurs without any of its causes. Hence, $A \wedge \neg D$ is minimally sufficient for C, and so for B. For this type structure, we obtain the true complex regularity:

$$(A \wedge \neg D) \vee C \vee E \leftrightarrow B. \tag{2}$$

A is an INUS condition of B. But an instantiation a of A should never count as a cause of an instantiation b of B—not even when A is co-instantiated with $\neg D$. In this case, a and b are merely joint effects of an occurring common cause c.

Baumgartner's insight is that $(A \wedge \neg D) \vee C \vee E$ is not minimally necessary for B because $C \vee E$ is still necessary for B. Indeed, B is only instantiated if C or E is. And $C \vee E$ is necessary for B in a minimal way: no disjunct can be removed without losing the necessity for B. $C \vee E$ is a minimally necessary disjunction of minimally sufficient conjunctions for B. In general, Baumgartner requires that each complex regularity must be a minimally necessary disjunction of minimally sufficient conjunctions for an effect. He calls such regularities 'non-redundant'.

The non-redundant regularities are relative to the set of considered variables. In the above structure, $(A \wedge \neg D) \vee E \leftrightarrow B$ is a non-redundant regularity relative to the variable set $\{D, E, A, B\}$. Extending the variable set by C, however, renders the regularity redundant. $(A \wedge \neg D) \vee E$ is not necessary for B any longer. B may be instantiated if A and D is, and E is not—namely when C is also instantiated. The non-redundant regularity of B after the extension is of course $C \vee E \leftrightarrow B$.

Baumgartner defines token causation in terms of type causation. He says, roughly, that C is a type cause of E iff C is a condition in a non-redundant regularity for E and remains so under any suitable extension of the variable set. An extension of the considered variables is *suitable* only if the additional variables do not introduce dependences among the variables that are stronger than causation, such as logical or mereological relations, supervenience, or grounding.

The assumption of multiple type causes ensures that all the non-redundant regularities are directed. For assume there is a simplistic regularity like $C \leftrightarrow E$, which has only one type cause for a type effect relative to some variable set. This simplistic regularity is non-directed: E is minimally sufficient for C, and C is minimally sufficient for E. However, by the assumption of multiple type causes, the variable set can either be suitably extended by another type cause C' of E or else by another type cause E' of C. The resulting non-redundant regularity, let's say $C \wedge C' \leftrightarrow E$, is directed: $\neg C$ and $\neg C'$ alone are minimally sufficient for $\neg E$, whereas $\neg E$ is only sufficient for their disjunction $\neg C \vee \neg C'$. This establishes the direction of non-redundant regularities and so the direction of type causation if the assumption of mutliple type causes is true (Baumgartner, 2013, pp. 94-8).

Equipped with his theory of type causation, Baumgartner defines token causation roughly as follows. A token event c is a cause of another e iff C is a type cause of E and there is an active path of direct non-redundant regularities from C to E. An active path of direct non-redundant regularities is a sequence $\langle C, D_1, ..., D_n, E \rangle$ of conditions, where each condition except E belongs to a direct minimally necessary disjunction of a minimally sufficient conjunction for its successor, and each condition except E is co-instantiated with all conditions of the respective minimally sufficient conjunction for its successor.

Baumgartner's regularity theory reduces causation to material implications and minimization procedures. No modal notions like counterfactuals are required. Moreover, Baumgartner's theory accounts well for many causal scenarios. His theory delivers the desired verdicts for overdetermination scenarios, preemption, as well as some short-circuits, and some switching scenarios. We think it therefore justified to say that Baumgartner gives the 'typical' regularity theory its present form.

We could explain our notion of law-like proposition in terms of Baumgartner's notion of direct non-redundant regularities: law-like propositions of a causal model in our framework are direct non-redundant regularities which are true of the respective causal scenario. This explanation would make our theory just as reductive as Baumgartner's and would place it within the tradition of 'typical' regularity theories (Andreas and Günther, forthcoming). Causation would be reduced to true propositions of particular fact and, optionally, facts about deviancy from norms. But we refrain from doing so because it remains unclear how a reductive theory can be adequately applied to causal models featuring simplistic regularities.

6.1.2 The Challenge of Applicability

How can Baumgartner's theory be applied to causal scenarios? Well, we model the scenario under consideration by a set of instantiated and non-instantiated variables and a set of regularities which remain non-redundant under any suitable extension of the variable set. Such a model does, however, not suffice to apply his theory without worries. He must and does, in addition, assume that the model of a causal scenario is complete. Otherwise his theory may come to the wrong verdicts about token causation, as we will show now.

Suppose the regularity $C \vee A \leftrightarrow E$ is true and non-redundant relative to the variable set $\{C, A, E\}$, and each variable is instantiated. Then the instantiation a of A is a cause of the instantiation e of E on Baumgartner's theory. Indeed, c and a look like overdetermining causes of e. But appearances may be deceptive. The actual scenario may be the preemption scenario depicted in Figure 3. The regularity is still true and non-redundant in this scenario relative to $\{C, A, D, B, E\}$: E is instantiated iff C or A is. But the instantiation of A which is preempted by the one of C should not count as a cause of e. This problem is quite general: our causal verdicts may very well change when we consider more variables—even if the relevant regularities remain non-redundant under the extension.

Baumgartner (2013, pp. 98-9) solves the problem by the *assumption of complete description*: the models of causal scenarios describe them completely. A complete description leaves no variables out and contains all direct non-redundant regularities. The regularity $C \vee A \leftrightarrow E$ does not describe the preemption scenario completely. It does not model that the efficacy of the instantiation of A is preempted by the simultaneous instantiation of C. Any complete description of the scenario, by contrast, does so. Take for example our causal model of the scenario, replace = by $\leftrightarrow$, and reverse the sides. We obtain the direct non-redundant regularities $C \leftrightarrow D, A \wedge \neg D \leftrightarrow B$, and $D \vee B \leftrightarrow E$. The actual events and absences are represented by $C, A, D, \neg B, E$. This complete description models why a is not a cause of e on Baumgartner's theory. There is no active path of direct non-redundant regularities from A to E. $A \wedge \neg C$ is minimally sufficient for B and B is minimally sufficient for E, but A is not co-instantiated with $\neg C$. The direct non-redundant regularities relative to $\{C, A, D, B, E\}$ completely describe the structure of the preemption scenario. Indeed, they entail the indirect regularity $C \vee A \leftrightarrow E$, which is thereby superfluous for a complete description. In sum, Baumgartner's theory is adequately applicable to causal scenarios only under the assumption of complete description.

The assumption of complete description entails that the model of a causal scenario contains *all* of the variables. As a consequence, the non-redundant regularities between the variables remain so under any suitable extension of the variable set—simply because there is none. Let us assume, for example, that the just discussed canonical model of the preemption scenario is complete. Then there are only the five variables $\{C, A, D, B, E\}$, and so this variable set cannot be extended. But this non-extendability in virtue of the assumption of complete description contradicts the assumption of multiple type causes. If the variable set cannot be extended, the type effect D can only have one type cause C. It follows that the direction of the simplistic regularity $C \leftrightarrow D$ in the scenario cannot be established under the assumption of complete description—at least not by Baumgartner's method of suitably extending the variable set. As a result, he is forced to assume the direction of simplistic regularities in his complete descriptions of causal scenarios.

Baumgartner faces a dilemma. The assumption of multiple type causes is essential to obtain the direction of the non-redundant regularities and so the direction of causation. His theory is not reductive if the assumption is given up. The assumption of complete description, on the other hand, is what allows us to adequately apply his theory to causal scenarios in the first place. If we give it up, we don't know what the *direct* non-redundant regularities are. And so we cannot check whether the paths of direct non-redundant regularities are active—a check his theory requires to determine whether this token is a cause of that. But as we have seen in the canonical preemption scenario, the two assumptions may well contradict each other. Indeed, they do so in any complete description which features a simplistic regularity. Hence, Baumgartner cannot make both assumptions—at least not in all causal scenarios.

In this paper, we treat Baumgartner's theory as prioritizing the assumption of complete description whenever it conflicts with the assumption of multiple type causes. Too many of the canonical causal scenarios discussed here and in the literature on token causation are modelled by simplistic regularities or corresponding structural equations. The point of the problem of unique causes is that it violates the assumption of multiple type causes: there is only a single type cause for the type effect. This being said, we are optimistic that there is a reductive regularity theory which can be applied to causal scenarios featuring simplistic regularities. One way to resolve the tension is to drop the assumption of complete description and to replace it by the assumption that the causal model under consideration is an abstraction of a causal model satisfying the assumption of multiple type causes. An abstraction of a causal model may abstract away from certain variables but not from others and the causal verdicts between the remaining variables must remain invariant. An investigation of the abstraction idea deserves its own paper.

The boulder scenario depicted in Figure 4 spells further trouble for Baumgartner's theory. To apply his theory, let us assume that its causal model corresponds to a complete description. Then the hiker's remaining unscathed is uncaused on his theory. The reason is that 'e's absence' has no type causes. For this to be seen, observe that the complete description corresponding to our causal model is empirically equivalent to the complete description featuring only the regularities $F \leftrightarrow B$, $F \leftrightarrow D$, and $\neg F \vee F \leftrightarrow \neg E$. In both complete descriptions, the respective sets of regularities allow for only two empirically possible situations: $\{F, B, D, \neg E\}$ and $\{\neg F, \neg B, \neg D, \neg E\}$.

For Baumgartner (2013, pp. 101-5) the empirically equivalent complete description shows that the regularity $B \wedge \neg D \leftrightarrow E$ is empirically redundant or "ungrounded". $\neg B \vee D$ is not a minimally necessary disjunction of minimally sufficient conjunctions for $\neg E$. The tautology $\neg B \vee B$ is a minimally sufficient 'conjunction' for $\neg E$, and so are the other tautologies $\neg D \vee D$ and $\neg F \vee F$. Indeed, the only minimally necessary disjunction of minimally sufficient conjunctions for $\neg E$ in the boulder scenario is some tautology. As a good result, the dislodged boulder does not count as a cause of the hiker's remaining unscathed. However, the hiker's ducking does also not count as a cause—which seems wrong.

We have learned that the underlying problem for Mackie's regularity theory—to establish the direction of causation—can be solved for complex enough scenarios, where each type effect has at least two type causes. Mackie only minimized the conjunctions or sets of actual conditions which are jointly sufficient for the effect. Baumgartner has seen that necessary conditions may also contain redundancies, and these redundancies must be minimized as well to avoid spurious regularities. Yet we have seen that

Baumgartner's theory is either reductive, or else adequately applicable to causal scenarios, but not both. Our regularity theory is 'typical' if we explain our law-like propositions in terms of non-redundant regularities. But then—without further ado—our theory would likewise face the challenge of applicability. Hence, we refrain from doing so for the time being.

Mackie (1974, pp. xiv & 85-6) gave up the ambitious quest for a reductive regularity theory of causation in the light of the problem of joint effects. Other authors departed as well from the the tradition of the 'typical' regularity theory of Hume and Mill over Mackie to Baumgartner. We will discuss their proposals next.

6.2 Non-Reductive Regularity Accounts

Wright (1985, 2011) builds on Hart and Honoré (1959/1985) to develop a regularity account similar to Mackie's (1965). The account roughly says a direct cause is an instantiated NESS condition for its effect: a cause is a *n*ecessary *e*lement of a *s*ufficient *s*et for the effect. Less roughly, a token event *c* is a direct cause of another *e* iff the condition *C* is a necessary element in a set of actual conditions that are jointly sufficient in a causal way for an instantiation of *E*. In many scenarios, *C* is a NESS condition for *E* iff *C* is at least an INUS condition for *E*. A NESS condition is a non-redundant part of a causally sufficient condition.

Unlike Mackie and like Strevens (2007), Wright (2011, pp. 289-90) employs a notion of causal sufficiency. A set of actual conditions is causally sufficient for an effect iff all antecedent conditions in a causal law are instantiated. A causal law specifies a minimal set of actual conditions that entails the immediate instantiation of some effect. "Immediate" means here that the effect occurs shortly after the instantiation of all antecedent conditions. Wright seems to use the direction of time to obtain the direction of causal laws, and thus the direction of causation. He writes as if he subscribes to the Humean dictum: causes must precede their effects in time.

This being said, Wright (2011, fn. 33) also writes:

> Interpreted in the usual manner, causal succession precludes

> temporally backward causation, through which events today change events in the past. However, the definition of causal succession in the text does not preclude such backward causation, which would occur if the present instantiation of the antecedent results in the immediately following instantiation of the consequent (paradoxically) in the past.

This reads paradoxical indeed. Pace Wright (2011, pp. 295-6), the directionality of the causal laws remains unexplained. He owes us an explanation why, for example, the true regularity $A \rightarrow E$ in the scenario of joint effects is not a causal law. After all, the regularity specifies a minimal set of actual conditions $\{A\}$ that entails the instantiation of the joint effect e a moment later. A similar point applies to the true regularity $A \rightarrow E$ in the preemption scenario. Given that the preempted condition A is instantiated, E will be instantiated a moment later—either because the genuine cause condition C is instantiated, or because it is not. Indeed, $\{A\}$ *is* a minimal set of actual conditions that entails the occurrence of e. So why is $\{A\}$—in both scenarios—not causally sufficient for e?

Wright (2011) gestures at Mill's difference method, and empirical observation and experimentation more generally. We observe in an experiment what happens after some manipulation in order to identify the effects of the manipulation. This seems to presuppose the Humean dictum of the temporal succession of cause and effect. Otherwise we cannot exclude that the manipulation caused a past event, which in turn caused the observed events. As we have just seen, Wright allows for backward causation: a cause may indeed obtain later in time than its effect. The Humean dictum is thereby jettisoned. And yet this dictum seems necessary to establish causal laws by observation and experimentation. The verdict stands: it remains unclear on Wright's account how the directionality of causal laws is determined.

Of course, Wright may rely on Baumgartner's non-redundant regularities as causal laws (on pain of inheriting the problem of applicability explained in Section 6.1.2). Without such an amendment, however, Wright's account does not account for the directionality of causal laws in terms of non-causal facts, and hence is not reductive. Strevens (2007), by contrast, acknowledges the non-reductive character of his regularity account: the primitive causal relations on the type level must somehow be determined by the physical laws.

Wright's account is transitive by stipulation. He says c is a direct cause of e iff c and e instantiate a causal law. The right-hand side means C is a necessary element in a set of actual conditions that is the complete antecedent of a causal law whose consequent is E. Finally, c is a cause of e iff there is a sequence of direct causes from c to e. A cause c is connected to its effect e by a sequence of instantiated causal laws.

Recall the preemption scenario. Under the restriction to the five variables C, A, D, B, E, there are four causal laws: $A \land \neg C \to B, C \to D, D \to E$, and $B \to E$. Consider the variation of the preemption scenario, where C is not instantiated, and so D is not, but A is instantiated and hence is B and E. In this scenario, our regularity theory says that a causes e via b, and the absence $\neg C$ does not cause e. We take this to be commonsensical. Wright's account, by contrast, wrongly says that the absence $\neg C$ is a cause of e. $\neg C$ is a necessary element in the set $\{A, \neg C\}$ of actual conditions that entails B by the causal law $A \land \neg C \to B$. And B is a necessary element in the set $\{B\}$ of actual conditions that entails E by the causal law $B \to E$.

Wright's account also leads to troublesome verdicts in scenarios that suggest that causation is not transitive. Recall the boulder scenario. The dislodged boulder causes the ducking of the hiker, which in turn causes the hiker's remaining unscathed. Wright's account says so. However, Wright's account must in virtue of its transitivity say that the dislodged boulder is a cause of the hiker's remaining unscathed. But this seems wrong.

A similar point applies to the simple switch. Flipper flips the switch f causing the train to travel down the right track r, instead of the left. The train travelling on the right track r causes the train to arrive e. F is an instantiated NESS condition for R, and R is an instantiated NESS condition for E. Wright's account is thus forced to say that the flipping of the switch is a cause of the train's arrival. This seems, again, wrong.

In the switch and boulder scenario, F is not a necessary element of a set of actual conditions jointly sufficient for E and $\neg E$, respectively. F is neither a NESS nor an INUS condition for E in the simple switch and $\neg E$ in the boulder scenario. Hence, Mackie's non-transitive theory comes to the desired verdicts.

Baumgartner's (2013) regularity theory is likewise not transitive. f in the simple switch scenario does not count as a cause of e. The reason is that F is no type-level cause of E: F can be removed from any set of conditions which are jointly sufficient for E without losing the set's sufficiency, and so F is no condition in any non-redundant regularity for E. In the confines of the scenario, the only minimally sufficient condition for E is the tautology $F \vee \neg F$. Similarly, as we have seen above, the falling boulder is no cause of the hiker's remaining unscathed on his theory.

This being said, Baumgartner's theory judges that the train travelling down the right track is not a cause of the train's arrival in the simple switch; and that the ducking of the hiker is not a cause of the hiker's remaining unscathed. The underlying reason is that the only minimally sufficient condition for the respective effect is a tautology, and so the effects are uncaused. Our theory, by contrast, delivers the desired verdicts.

6.2.1 Formalisations of the NESS Account

We have embedded Lewis's regularity theory into causal models and refined it. Others had the idea to embed Wright's (1985) NESS account into causal models. The idea surfaced first in Baldwin and Neufeld (2003, 2004). However, their account is not strictly speaking a NESS account, but rather a *de facto* account: an effect counterfactually depends on a genuine cause when holding certain events and absences fixed by intervention. Holding this and that fixed, the effect would not have obtained if the cause had not obtained. Wright (2011, pp. 287&304) by contrast, stays clear of counterfactuals and aims for a "factual" account. This is, in part, why Beckers (2021b, p. 6215) writes that Baldwin and Neufeld's account "is inconsistent with Wright's views of the NESS definition."

Halpern (2008, pp. 205-7) aims to formalize Wright's (1985) NESS condition in Halpern and Pearl's (2005) framework of causal models. Roughly, C is a Halpern-NESS condition of E in a causal model if C belongs to some set $\mathcal{S}$

of actual events and absences such that $\mathcal{S}$ is strongly sufficient for E in the causal model, and $\mathcal{S} \setminus \{C\}$ is not. In an attempt to clarify Wright's notion of causal sufficiency, he says a set $\mathcal{S}$ of events and absences is strongly sufficient for E in a causal model if $\mathcal{S}$ remains sufficient for E when adding any actual events and absences to it. A set $\mathcal{S}$ of events and absences is sufficient for an effect E in a causal model if setting $\mathcal{S}$ by intervention entails E in the resulting submodels across different 'contexts', including non-actual ones.

A Halpern-NESS condition is, however, inadequate as a formalisation of a NESS condition. As we have observed above, F is not a NESS condition for E in the simple switch and $\neg E$ in the boulder scenario, but it is a Halpern-NESS condition for each. And so the flipping of the switch counts as a cause of the train's arrival and the dislodged boulder counts as a cause of the hiker's remaining unscathed on Halpern's (2008) NESS test. Another counterexample, where a genuine NESS condition does not count as a Halpern-NESS condition may be found in Beckers (2021b, p. 6214).

Beckers (2021b, pp. 6213-4) also provides another formalisation of Wright's (2011) NESS account in Halpern and Pearl's framework of causal models. He represents Wright's non-reductive causal laws by likewise non-reductive structural equations. This allows to define a notion of causal sufficiency as sufficiency in causal models. The resulting NESS account is stipulated to be transitive. And so it inherits the problems of the original NESS account in the boulder and switch scenarios. Moreover, it counts the absence $\neg C$ a cause of e in the variation of preemption discussed above—which seems wrong to us. Beckers's formalisation of the NESS account resembles the original indeed.

Moreover, Beckers (2021b, p. 6216) proposes an "improvement". He marries his NESS account with a counterfactual condition: if c is a cause of e, then, had c not obtained, its absence would not have been a cause of e. Sartorio (2006, pp. 73-5) motivates this principle by switching scenarios. According to Sartorio's principle alone, flipping the switch cannot be a cause of the train's arrival in the simple switch because not flipping the switch would be a cause of the train's arrival as well.

Beckers's counterfactual NESS account defines causation in terms of his

NESS causation coupled with a path-specific version of Sartorio's principle. c is a CNESS cause of e if c is a NESS cause of e along some path p in the causal model M and $\neg C$ is not a NESS cause of e along any subpath of p in the causal submodel of M after intervening by $\neg C$. Notwithstanding Sartorio's motivation, flipping the switch is a CNESS cause of the train's arrival. Flipping the switch is a NESS cause of the train's arrival via its travelling on the right track. And not flipping the switch would not be a NESS cause of the train's arrival via its travelling on the right track. The train's merely possible path along the left track is quite literally no subpath of the actual path to its destination. A similar argument shows that the dislodgement of the boulder is a CNESS cause of the hiker's remaining unscathed.

Beckers (2021b, pp. 6210&6216) says his CNESS account is a "nice" and "natural" compromise of a regularity account and a counterfactual one. He does, however, not explain why such a compromise is desirable.

The CNESS account is a simpler version of Beckers's (2021a) definition of causation. He claims that the latter definition is "a formal expression of the NESS intuition" (p. 1352). But he employs a counterfactual notion of necessity instead of a notion of non-redundancy: when testing for causation, the putative cause is replaced by a non-actual event or absence rather than simply removed from the minimal set of actual conditions sufficient for the effect. He roughly defines causation to be the transitive closure of direct sufficiency coupled with a network-specific version of Sartorio's principles. This is not a formal expression of Wright's NESS account, as Beckers admits (p. 1342, fn. 1). He also acknowledges that the explicit statement of his favourite definition "looks even more complicated than" Halpern and Pearl's (2005) de facto definition (p. 1354). Except for one of the many examples in the 2005 paper, the two definitions come to the same verdicts (p. 1358). Moreover, the definition agrees with the CNESS account on the verdicts in the simple switch and boulder scenarios. So why should we settle for it?

Beckers (2021a, pp. 1361-3) argues that his favourite definition delivers "consistent (and intuitive) answers" to a series of closely related scenarios—unlike many other accounts of causation, including the de facto definitions of Halpern and Pearl (2005) and Halpern (2015). We leave it to the reader to verify that our regularity theory delivers the results Beckers desires in the series of scenarios. One of his selling points supports our theory as well.

6.3 Counterfactual Accounts

Our regularity theory does not rely on any condition of counterfactual dependence. It does *not* ask what would have happened, had the putative cause not obtained. Thereby our theory does not rely on counterfactual dependence, de facto dependence, or Sartorio's principle—unlike the accounts of Beckers and Vennekens (2017, 2018) and Beckers (2021b,a) for example. Our regularity theory is not counterfactual.

In this section, we briefly explain counterfactual accounts of causation and discuss a switching scenario proposed by Halpern and Hitchcock (2010). Finally, we say a few words on Gallow's (2021) account—one of the leading counterfactual accounts at the moment.

The starting point of counterfactual accounts of causation is that counterfactual dependence between distinct occurring events is sufficient for causation. The simple counterfactual account elevates counterfactual dependence between actual events and absences to a necessary and sufficient condition for causation. The token event c is a cause of a distint token event e iff c and e are actual, and had c not been actual, e would not have been actual. Notably, the simple counterfactual account solves the simple switch and the boulder scenario. Had the switch not been flipped, the train would have arrived at its destination anyways. Had the boulder not been dislodged, the hiker still would have remained unscathed. The flipping of the switch and the dislodgement of the boulder do not make a difference to the train's arrival and the hiker's remaining untouched, respectively. Moreover, on a non-backtracking interpretation of counterfactuals, the train travelling on the right track is a cause of the train's arrival, and the ducking is a cause of the hiker's remaining unscathed.

As is well-known, however, the simple counterfactual account has troubles with scenarios of preemption. Had the genuine cause c not occurred, the effect e would still have occurred—due to the backup cause a. Hence, the genuine cause c does not count as a cause. In response, Lewis (1973) says that causation is the transitive closure of non-backtracking counterfactual dependence between actual events and absences. This solves certain preemption scenarios, but not others. Unfortunately, it also makes flipping the switch a cause of the train's arrival. There is a chain of true non-backtracking counterfactuals running from flipping the switch over the train's travelling on the right tracks to its arrival at the destination. The dislodgement of the boulder likewise counts as a cause of the hiker's remaining unscathed.

There are plenty de facto accounts of causation using causal models (Hitchcock, 2001; Woodward, 2003; Halpern and Pearl, 2005; Halpern, 2015). For the simple switch, they have all in common that flipping the switch counts as a cause of the train's arrival (Blinded, forthcoming). For the train's arrival counterfactually depends on flipping the switch when holding fixed by intervention that the train does not travel on the left tracks. And similarly for the boulder scenario.

This being said, Halpern (2016, pp. 79-81&90-1) shows how the definitions of Halpern and Pearl (2005) and Halpern (2015) can be amended by a condition of normality so that they solve the simple switch. Roughly, causation is then understood as de facto dependence witnessed by a possible world which is at least as normal as the actual one. The idea is that the non-actual world, where the train does not travel on the left track even though the switch has not been flipped, is less normal than the actual world. Hence, there is no possible world at least as normal as the actual witnessing that the train's arrival de facto depends on the flipping.

Another resort for causal modellers when their accounts deliver an undesired result is to say that the causal model employed to represent the causal scenario is inappropriate. Halpern and Hitchcock (2010, Sec. 4.3) argue that values of different variables in a causal model must be logically independent, and further that the variables R and L in the simple switch are

> arguably not independent; the train cannot be on both tracks at once. If we want to model the possibility of one track or another being blocked, we should use, instead of [L and R],

> variables LB and RB, which indicate whether the left track or right track, respectively, are blocked. This allows us to represent all the relevant possibilities without running into independence problems.

We disagree: the variables R and L are not logically dependent. As Beckers and Vennekens (2017, p. 14) put it, "it is a matter of physics, not logic, that a train can only occupy a single track at any given moment."

Halpern (2016, pp. 38-9) proposes the modified switch scenario, where the tracks are unblocked but might be blocked, in an attempt to save the verdict that flipping the switch is on Halpern's (2015) definition not a cause of the train's arrival. Here is his causal model:

$E = (F \wedge \neg RB) \vee (\neg F \wedge \neg LB)$
$F, \neg RB, \neg LB, E$

Halpern (2016, pp. 38) says "it seems strange to call flipping the switch a cause of the train arriving when in fact both tracks are unblocked." Still, the definition of Halpern and Pearl (2005) says so. And the one of Halpern (2015) counts the flipping as 'part of' the cause $\{f, \neg lb\}$, where parts of causes correspond to "what we think of as causes" (Halpern, 2016, p. 25). The definitions amended by a condition of normality overcome the problem if the non-actual world, where the left track is blocked, is less normal than the actual world. For then, there is no de facto dependence of e on f witnessed by a possible world which is at least as normal as the actual one.

Our regularity theory without the deviancy condition likewise says that the flipping of the switch f is a cause of the train's arrival e. Conditions (i)-(iv) are satisfied for $\mathcal{L}' = \mathcal{L}$ and $\mathcal{F}' = \{\neg RB\}$. Indeed, flipping the switch is a member of a maximised minimal set $\{F, \neg RB\}$ of actual conditions which, in the presence of the law-like proposition, entails the effect in a forward-directed way. f is also an insufficient but non-redundant part of an instantiated sufficient condition for e. F is an INUS condition of E in Halpern's switch. In the simple switch, by contrast, flipping the switch is redundant, which is arguably a feature of typical switching scenarios.

Our regularity theory amended by the condition (v) of deviancy, however, says that the flipping of the switch f is not a cause of the train's arrival e. For this to be seen, note that $\neg LB$ is in $\mathcal{F} \setminus \mathcal{F}'$ and the variable LB is neither a descendant nor an ancestor of F, and yet $\neg LB$ is less deviant than LB. Hence, f is not a cause of e, as desired in Halpern's switch.

We have seen that switching scenarios pose problems for many accounts of causation. It is thus not surprising that their representation is controversial. Our regularity theory without deviancy condition delivers the desired results for the "basic" switch discussed by Paul and Hall (2013, p. 232). Amended by the deviancy condition, our theory also delivers the desired results for the more "realistic" switches discussed by Hitchcock (2009, p. 395-6). The amendment by the deviancy condition may well be worth it.

If we choose the optional deviancy condition, counterfactual dependence is clearly not sufficient for causation on our theory. For this to be seen, reconsider the basic structure of omissions. The absence $\neg F$ does not prevent e from occurring. But had f occurred, e would not have occurred. Still, Putin's failure to water the flowers is not a cause of their death. For Putin's omission to water the flowers is not deviant—he did not promise to water them.

6.3.1 Gallow's Account

Gallow (2021) offers perhaps the most sophisticated counterfactual account of causation. On closer inspection, he actually offers several closely related accounts. One is guided by the idea that a cause must transmit deviancy via an active causal network to its effect. Roughly, each member of a set $\mathcal{C}$ of particular propositions, or variable assignments, is a cause of E in a causal model M iff there is a minimal causal network in M leading from $\mathcal{C}$ to E, and the propositions in $\mathcal{C} \cup \{E\}$ are more deviant than their respective negations (p. 83). A network consists of directed paths, which start from some $C \in \mathcal{C}$ and end up in E. In a causal network, the value of each variable not in $\mathcal{C}$ counterfactually depends on certain values of its parent variables. Such dependences are called *local*.

This counterfactual account can handle an impressive set of scenarios including some switches, but it has troubles in the simple switch. There is a minimal causal network leading from flipping the switch $\{F\}$ to the train's arrival E: $F \rightarrow R \rightarrow E$. Within this causal network, E counterfactually depends on R, and R counterfactually depends on F. Moreover, the proposition F departing from the minimal network to $\neg L$ and the return proposition E are both more deviant than their negations. Hence, flipping the switch counts as a cause of the train's arrival on Gallow's account.

Gallow (2021, p. 87) himself observes a consequence of his deviancy requirement: "default, inertial states can be neither causes nor effects." This means that preventers do not count as causes in simple prevention scenarios. Assassin poisons target's coffee. Bodyguard prevents target's death by putting antidote in her coffee. It seems that bodyguard's putting in the antidote causes target's default survival. But the present account must deny causation here and likewise for omissions which are supposedly causal.

The problem with genuine prevention cases and omissions motivates Gallow (2021, p. 88) to mention three variants of the above theory. These variants agree that, for C to be a cause of E, there must be a minimal causal network in M leading from C to E. They differ in what actual values of the cause and effect variables must be deviant. There are three options: (i) causes must be deviant, but not effects; (ii) effects must be deviant, but not causes; (iii) neither causes nor effects must be deviant. The variants no longer transmit deviancy from cause to effect.

Gallow doesn't say which of the constraints on deviancy should be preferred. We recommend variant (i): causes must be deviant, but not effects. With this constraint in place, it is easy to show that Gallow's theory discriminates between bogus and genuine simple preventions in the same way our theory does. Likewise, the discrimination between supposedly causal and presumably non-causal omissions is not a problem any more for Gallow's theory. We merely have to declare that a violation of a norm is more deviant than conforming to it. If a neighbour fails to water the plants despite promising to do so, this is then recognized as a cause of the death of the plants. Putin's not watering these plants is not as long as he doesn't have an obligation to do so.

Finally, recall the boulder scenario. The hiker's remaining unscathed by the dislodged boulder is default. If effects are admitted to have non-deviant values, Gallow's account runs into a problem: it says that the dislodgement of the boulder and its rolling toward the hiker are joint causes of the hiker's remaining unscathed. There is a minimal causal network leading from $\{F, B\}$ to $\neg E$: $F \rightarrow D \rightarrow E \leftarrow B$. To verify that there is such a network, we need to specify contrasts for the values of F, B, and D. Since we are free to assign, for all non-effects, a contrast which does not differ from the actual value of the variable, we can choose the following contrasts: F and D are false, while B is true. Then, it holds for both D and E that their value locally depends on the values of their parents, and so $F \rightarrow D \rightarrow E \leftarrow B$ is a causal network. Minimality is easy to show for this network. Hence, the dislodged boulder is a cause of the hiker's remaining unscathed on variant (i) of Gallow's account—a joint cause with the boulder's rolling toward the hiker. This is, of course, an unfortunate verdict.

We suggest two solutions for the problem that the dislodged boulder counts as a cause on variant (i) of Gallow's account. First, we may require that all members of the set C of presumed causes have contrasts which differ from their actual values. Second, we may require that any assignment of contrasts to a variable's parents must satisfy all the structural equations of the causal model. This being said, these solutions may of course lead to troubles in other causal scenarios.

7 Conclusion

We have refined Lewis's regularity theory of causation by embedding it into a framework of causal models. Our theory says that causation is *forward-directed inferability along lawful paths*. It solves the problems that speak decisively against Lewis's regularity theory: the problems of unique causes, joint effects, and preemption. It can also handle causal scenarios which suggest that causation is not transitive, like the simple switch and boulder scenarios.

We have offered an optional condition of deviancy. Our theory amended by this condition can address the problem of isomorphic models. We have shown that our amended theory delivers the desired verdicts for the bogus prevention scenario, even though this scenario is isomorphic to a scenario of overdetermination. Moreover, our theory says that deviant omissions are genuine causes while non-deviant omissions are not. Perhaps the deviancy condition earns its keep, as it also helps our theory to account for many switching scenarios.

We have argued that the present form of the 'typical' regularity theory is given by Baumgartner's (2013). His theory reduces causation to material implications and minimization procedures. He thereby proposes a theory of causation free of counterfactuals, any other modal notion, or epistemic notions (Andreas and Günther, 2019, 2020; Andreas and Günther, 2021a). And it is notable that his theory accounts well for many causal scenarios, including overdetermination, preemption, as well as some switching scenarios, and some short-circuits. Only recently counterfactual theories of causation have been able to account for these scenarios (Andreas and Günther, 2021b; Gallow, 2021).

Baumgartner (2013, p. 106) prefers not to amend his regularity theory by a notion of deviancy or typicality. He points to the intuition that causation is an entirely objective matter that is independent of contexts and norms. This being said, he outlines how his theory could be amended by a notion of deviancy. He can thereby secure the verdict in the bogus prevention scenario that bodyguard's putting in the antidote is not a cause of target's survival. However, he must still say that assassin's refraining to poison target's coffee is a cause of target's survival. But this goes against common sense: the typical absence of poison does not cause target's survival.

Moreover, Baumgartner's theory, like Andreas and Günther's (2021b), has troubles with the simple switch and boulder scenarios. The theories say, against common sense, that the train travelling down the right track is not a cause of the train's arrival in the simple switch, and that the ducking of the hiker is not a cause of the hiker's remaining unscathed. We have also pointed out that Gallow's (2021) counterfactual accounts of causation have troubles with these examples.

We have discussed that we could rely on Baumgartner's non-redundant regularities to save the reductivity of our theory. But then, our theory would—just like Baumgartner's—face the challenge of applicability: it would not be adequately applicable to many of the causal scenarios discussed in the literature. We hope to overcome this challenge in future work.

Our theory amended by the condition of deviancy is in a way still incomplete. We haven't said much on what norms are and when events deviate from norms in a given scenario. In future work, our theory should be amended by a theory of what norms are. We can then also address the question whether or not norms can be reduced to propositions of particular matter of fact. For now it should suffice to say that our theory emerges as a competitor to the most advanced regularity and counterfactual accounts of causation.

References

Andreas, H., Armgardt, M., and Günther, M. (2022). Counterfactuals for Causal Responsibility in Legal Contexts. *Artificial Intelligence and Law* .

Andreas, H. and Günther, M. (2019). "On the Ramsey Test Analysis of 'Because'". *Erkenntnis* **84**(6): 1229–1262.

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

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

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

Andreas, H. and Günther, M. (forthcoming). A Regularity Theory of Causation. *Pacific Philosophical Quarterly* .

Andreas, H. and Günther, M. (2021). Regularity and Inferential Theories of Causation. In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta, Metaphysics Research Lab, Stanford University. Fall 2021 edn.

Baldwin, R. A. and Neufeld, E. (2003). The Structure Model Interpretation of Wright's NESS Test. In *Advances in Artificial Intelligence*, edited by Y. Xiang and B. Chaib-draa. Berlin, Heidelberg: Springer Berlin Heidelberg, 9–23.

Baldwin, R. A. and Neufeld, E. (2004). The Structural Model Interpretation of the NESS Test. In *Advances in Artificial Intelligence*, edited by A. Y. Tawfik and S. D. Goodwin. Berlin, Heidelberg: Springer Berlin Heidelberg, 292–307.

Baumgartner, M. (2013). A Regularity Theoretic Approach to Actual Causation. *Erkenntnis* **78**(1): 85–109.

Beckers, S. (2021a). Causal Sufficiency and Actual Causation. *Journal of Philosophical Logic* **50**(6): 1341–1374.

Beckers, S. (2021b). The Counterfactual NESS Definition of Causation. *Proceedings of the AAAI Conference on Artificial Intelligence* **35**(7): 6210–6217.

Beckers, S. and Vennekens, J. (2017). The Transitivity and Asymmetry of Actual Causation. *Ergo: An Open Access Journal of Philosophy* **4**: 1–27.

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

Beebee, H. (2004). Causing and Nothingness. In *Causation and Counterfactuals*, edited by L. A. Paul, E. J. Hall, and J. Collins, Cambridge, MA, USA: MIT Press. 291–308.

Blinded (forthcoming). Actual Causation. *dialectica* .

Dowe, P. (2000). *Physical Causation*. Cambridge Studies in Probability, Induction and Decision Theory, Cambridge University Press.

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

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

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

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

Halpern, J. Y. (2008). Defaults and Normality in Causal Structures. In *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning*. KR'08, AAAI Press, 198–208.

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

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

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

Hart, H. L. A. and Honoré, T. (1959/1985). *Causation in the Law*. Oxford University Press, 2nd edn.

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

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

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

Hitchcock, C. (2009). Structural Equations and Causation: Six Counterexamples. *Philosophical Studies* **144**(3): 391–401.

Hume, D. (1748/1975). In *An Enquiry Concerning Human Understanding*, edited by L. A. Selby-Bigge and P. H. Nidditch, Oxford: Clarendon. 3rd edn.

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

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

Mackie, J. (1974). *The Cement of the Universe: A Study of Causation*. Clarendon library of logic and philosophy, Clarendon Press.

Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly* **2**(4): 245–264.

McDermott, M. (1995). Redundant Causation. *The British Journal for the Philosophy of Science* **46**(4): 523–544. doi:10.1093/bjps/46.4.523.

McGrath, S. (2005). Causation By Omission: A Dilemma. *Philosophical Studies* **123**(1-2): 125–148.

Mill, J. S. (1843/2011). *A System of Logic*, vol. 1 of *Cambridge Library Collection - Philosophy*. Cambridge University Press.

Paul, L. A. (2000). Aspect Causation. *Journal of Philosophy* **97**(4): 235–256.

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

Pearl, J. (2009). *Causality: Models, Reasoning and Inference*. New York, NY, USA: Cambridge University Press, 2nd edn.

Psillos, S. (2002). *Causation and Explanation*. Routledge.

Sartorio, C. (2005). Causes as Difference-Makers. *Philosophical Studies* **123**(1/2): 71–96.

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

Schaffer, J. (2005). Contrastive Causation. *The Philosophical Review* **114**(3): 327–358.

Strevens, M. (2007). Mackie Remixed. In *Causation and Explanation*, edited by J. K. Campbell, M. O'Rourke, and H. S. Silverstein, MIT Press. 4–93.

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

Wright, R. W. (1985). Causation in Tort Law. *California Law Review* **73**(6): 1735–1828.

Wright, R. W. (2011). The NESS Account of Natural Causation: A Response to Criticisms. In *Perspectives on Causation*, edited by R. Goldberg, Hart Publishing, chap. 14.

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