Counterfactual theories of causation and the problem of large causes

Jens Harbecke[1] ⬤

Abstract As is well-known, David Lewis' counterfactual theory of causation is subject to serious counterexamples in 'exceptional' cases. What has not received due attention in the literature so far is that Lewis' theory fails to provide necessary and sufficient conditions for causation in 'ordinary' cases, too. In particular, the theory suffers from the 'problem of large causes'. It is argued that this problem may be fixed by imposing a minimization constraint, whilst this solution brings along substantial costs as well. In particular, a precise formulation of minimization requires defining an 'essential part of an event' and/or an 'essential subevent'. Although the possibility of such a definition is ultimately left open, some doubts are raised on whether the counterfactualists' resources are fit for this purpose, and whether the challenge can be met without substantially departing from Lewis' intention, which was to provide a reductive account of causation.

Keywords Counterfactual causation · Theories of causation · David Lewis · Overdetermination · Proportionality · John L. Mackie · Minimization

---

✉ Jens Harbecke
jens.harbecke@uni-wh.de;
http://www.jensharbecke.com

[1] Witten/Herdecke University, Department of Philosophy, Politics, and Economics, ISIC - The PPE Institute for Social and Institutional Change, Alfred-Herrhausen-Str. 50, Witten 58448, Germany

1 Introduction

David Hume famously defined a cause as...

> ...an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed. (1748, sec. 7)

The first part of this quote, according to which like "objects" are always followed by like "objects", is commonly interpreted as expressing the core idea of a regularity theory of causation. The second part, introducing a modal dependence of a single "object" upon another one, is commonly considered as history's first explicit formulation of a counterfactual theory of causation.[1] Hume himself did not spell out this theory further, and the regularity theory remained at the center of research on Humean theories of causation until deep into the twentieth century. In particular, systematic elaborations of counterfactual theories of causation as the one of Lewis (1973a) became available only after a comprehensive analysis of counterfactual statements had been developed (cf. Lewis 1973b). However, as Lewis' seminal work began to reach an ever-growing audience, counterfactual theories gradually started to dominate the debate on causation within a Humean ontology.

The beauty of Lewis' approach lies in its simplicity and systematicity. A sufficiently rich modal ontology[2] and a similarity relation between worlds is all that is required for a counterfactual analysis of the notion of cause. The downside of the author's approach lies in several intricate problems, some of which still have not been fully accommodated. For example, it has been pointed out that Lewis' attempt to build a time asymmetry into his theory is unsuccessful (cf. Elga 2001; Frisch 2005, ch. 8).[3] Further, it has been noticed that counterfactual dependence is not transitive whereas Lewis claimed that causation is transitive (cf. Lewis 1973a, 563; Lewis 2000, 194/195; Mackie 1980). Finally, in cases of late preemption, double preemption, symmetric overdetermination, and trumping Lewis' analysis fails to identify the correct causes (cf. Schaffer 2000; Hall 2007, 119; Horwich 1993, McDermott 1995; Menzies 2009; and Lewis 1986e).

This paper discusses a further challenge for Lewis' theory that has not received due attention in the literature so far. It can be characterized as the 'problem of large causes'. The problem is a result of the fact that Lewis' original analysis sometimes declares as causes events that are radically out of scale from the point of view of the effect. It is argued that the problem can be solved if Lewis' theory is amended by a minimization constraint. However, a precise formulation of minimization requires defining an 'essential part of an event' and/or an 'essential subevent' without

---

[1] But see Hüttemann (2013, 38–39) for a critical note on this interpretation.

[2] But note that Lewis' theory of counterfactuals does not as such entail his modal realism. Counterfactuals require truth values, but whether that in turn requires the reality of possible worlds (or other possibilia) is an open question. I thank an anonymous referee for bringing this point to my attention.

[3] For more discussions of time asymmetry in a Humean setting, see also Loewer (1996, 2007), Albert (2012), and Frisch (2011).

building on prior causal knowledge. Although the possibility of such a definition is ultimately left open, some doubts are raised on whether the counterfactualists' resources are fit for this purpose, and whether the challenge can be met without substantially departing from Lewis' intention, which was to provide a reductive account of causation.

The investigation proceeds as follows. Section 2 reproduces the main ideas of Lewis' classical counterfactual theory of causation. Sections 3 and 4 then show that Lewis' theory faces some challenges even with respect to the class of "ordinary" cases for which it was primarily designed. Section 5 contains some proposals as to how the theory would have to be amended to avoid the explicated problems, and it discusses the potential limits of these proposals. Section 6 summarizes the paper and makes some concluding remarks on the significance of the established claims for the general project of developing a satisfactory Humean theory of causation.

2 Counterfactual causation

As mentioned above, Lewis uses Hume's idea of a counterfactual dependence between events as a starting point for his counterfactual theory of (deterministic) causation. He first defines the notion as follows (from now on the symbols c, c', c'', ..., c^*, c^{**}, ..., c_1, c_2, ..., e, e_1, e_2, ...etc. are used to refer to events; $O(c)$, $O(c')$, $O(c'')$, ..., $O(c^*)$, $O(c^{**})$, ...$O(e)$,...etc. stand for the propositions that the events occur; γ, ϕ, ψ, ϕ', ϕ'', ...stand for variables quantifying over the domain of events):

Counterfactual dependence: $O(\psi)$ counterfactually depends on $O(\phi)$ if, and only if, $O(\phi)\square\rightarrow O(\psi)$ and $\neg O(\phi)\square\rightarrow \neg O(\psi)$ (Lewis 1973a, 561/62)

The symbol '$\square\rightarrow$' is intended by Lewis as a counterfactual conditional operator roughly corresponding to the English subjunctive conditional schema "if ... had been the case, then ... had been the case" (cf. Lewis 1973b, c). Lewis then explicates the semantics for such conditionals in terms of a possible worlds space. According to this idea, a counterfactual $O(\phi)\square\rightarrow O(\psi)$ is true at a world w if, and only if...

...either (1) there are no possible $O(\phi)$-worlds (in which case $O(\phi)\square\rightarrow O(\psi)$ is *vacuous*), or (2) some $O(\phi)$-world where $O(\psi)$ holds is closer (to w) than is any $O(\phi)$-world where $O(\psi)$ does not hold. (Lewis 1973a, 560)

Below, we will refer to this definition as the *Lewis Analysis* of counterfactuals. The relation "closer to w" is based on the assumption that all possible worlds are weakly ordered with respect to similarity in particular fact and laws of nature (for more details, cf. Sect. 3). Note that, according to the Lewisian analysis for counterfactuals, when ϕ and ψ are actual events, the conditional $O(\phi)\square\rightarrow O(\psi)$ is trivially true and whether or not $O(\phi)$ counterfactually depends on $O(\phi)$ is wholly determined by the truth or falsity of the conditional $\neg O(\phi)\square\rightarrow \neg O(\psi)$.

In the final step of his proposal, Lewis argues that counterfactual dependence implies causal dependence, and that "[c]ausal dependence among actual events

implies causation." (Lewis 1973a, 563). In short, he proposes the following sufficient condition for causation (for two non-overlapping events ϕ and ψ):

Causation$_{Lewis\text{-}R}$:   If $O(\psi)$ counterfactually depends on $O(\phi)$, then ϕ causes ψ.

Within the debate about Lewis' theory it became soon clear that **Causation**$_{Lewis\text{-}R}$ alone does not do the trick. Suppose that an event c causes both e_1 and e_2, but e_2 occurs later than e_1. In this case it seems that e_2 would not have occurred if e_1 had not occurred, because if e_1 would not have occurred, then c would not have occurred either. Under **Causation**$_{Lewis\text{-}R}$, e_1 causes e_2. But, according to the hypothesis, e_1 does not cause e_2. Hence, **Causation**$_{Lewis\text{-}R}$ must be false.

Lewis responded that what he calls the "standard resolution" of counterfactuals avoids this problem. This notion presupposes that "...if the present were different the past would be the same, but the same past causes would fail somehow to cause the same present effects." (Lewis 1986b, 34) Hence, back-tracking counterfactuals such as the discussed $\neg O(e_1) \square\!\!\rightarrow \neg O(c)$ come out false − unless interpreted in a peculiar and non-standard way. Hence, **Causation**$_{Lewis\text{-}R}$ remains true – or so it is still widely held (cf. McDermott 1995, 523 and Kroedel 2008, 126 for two examples generally embracing **Causation**$_{Lewis\text{-}R}$).

Did Lewis hold that causal, and hence counterfactual, dependence is also necessary for causation? The answer is "No!":

> Causation must always be transitive; causal dependence may not be; so there can be causation without causal dependence. Let c, d, and e be three actual events such that d would not have occurred without c and e would not have occurred without d. Then c is a cause of e even if e would still have occurred (otherwise caused) without c. (Lewis 1973a, 563)

On the other hand, Lewis pointed out that causal, and hence counterfactual, dependence can be made a transitive relation in the usual way. He uses the notion of causal chain to make the idea precise. A causal chain can be defined as follows:

Causal chain: If ψ and ϕ are connected by a causal chain, then there exists a sequence of events γ_1, γ_2, ..., γ_{n-1}, γ_n, such that $O(\psi)$ counterfactually depends on $O(\gamma_n)$, which counterfactually depends on $O(\gamma_{n-1})$, which ..., ..., which counterfactually depends on $O(\gamma_1)$, which counterfactually depends on $O(\phi)$.

With the help of this concept, Lewis can state a reductive account of causation:

Causation$_{Lewis\text{-}R}$: ϕ causes ψ if, and only if, "there exists a causal chain leading from [ϕ] to [ψ]." (Lewis 1973a, 563).

And since, according to **Causal Chain**, a causal chain can be reduced to a chain of counterfactual dependence, the theory can be stated without the term "causal" appearing on the right-hand side of the conditional, which makes its reductive character apparent:

Causation$_{Lewis}$: ϕ causes ψ if, and only if, there exists chain of counterfactual dependence leading from ϕ to ψ.[4]

As Lewis already acknowledged, there are certain "special" (1986b, 43) or "exceptional" (1986c, 242) cases that falsify this general assumption.[5] An example discussed by Lewis already in his (1973a, 567) is that of early preemptive causes. Forty years of debate of Lewis' proposal have brought to light a whole variety of further potential "exceptional" circumstances that are not easily accommodated within the original approach.[6]

Despite these complications of Lewis' original account, it has been widely agreed that **Causation** $_{Lewis}$ is adequate for "ordinary" circumstances in which there is only a non-preempted single non-redundant cause for a given effect. And even if the mentioned "exceptional" cases occur in the actual world more often than anticipated by Lewis, this appears to be an important insight in itself.[7]

The central aim of Sects. 3 and 4 will be to challenge this common perception by showing that, even relative to many "ordinary" circumstances, **Causation** $_{Lewis}$ is inadequate because incomplete. On the basis of this finding, it is argued in Sect. 5 that even the original counterfactual theorist who limits her analysis to "ordinary" cases is committed to a more complex theory than **Causation** $_{Lewis}$.

3 Overly large causes

Counterfactuals are variably strict conditionals as their antecedents cannot always be supplemented *salva veritate* by further conjuncts (cf. Lewis 1973b, 13–19). Notwithstanding, it is possible to show that, if the event ϕ' chosen for the supplementation is both sufficiently remote from ϕ and large enough, the supplementation *does* always preserve the truth of the counterfactual.

To illustrate this claim, consider the model of Fig. 1. In this picture, @ is the actual world; the circles around @ depict sets of worlds ordered according to their similarity to @. The dotted lines indicate various partitions of the total set of possible worlds relative to the occurrence, or non-occurrence, of particular events in those worlds. For example, the dotted line closest to @ indicates that the worlds to

---

[4] In Lewis' own words: "Causation is the ancestral of causal dependence: event c causes event e iff either e depends on c, or e depends on an intermediate event d which in turn depends on c, or.... Causation without direct causal dependence is exceptional, but it occurs in cases of causal preemption." (Lewis 1986c, 242). The abbreviation for "if, and only if" clearly states a reductive approach.

[5] Below, we routinely distinguish "exceptional" from "ordinary" cases of actual causation, where a case is "ordinary" if it involves a single non-preempted and non-redundant cause for a given effect.

[6] Among these are cases of symmetric overdetermination (Lewis 1973a, fn. 12), late preemption (Lewis 1986e), trumping preemption (Schaffer 2000), and double prevention (Hall 2007, 119; cf. also Horwich 1993, McDermott 1995, Menzies 2009).

[7] It should be noted that, even for "ordinary" cases, the simple account must be supplemented by a "ban on replacement" (Bennett 2003, fn. 21) with respect to the worlds used for evaluating counterfactuals if counterfactual dependence should be necessary for causation (cf. also Lewis 2000, 190 and Lewis 1986e, 211).
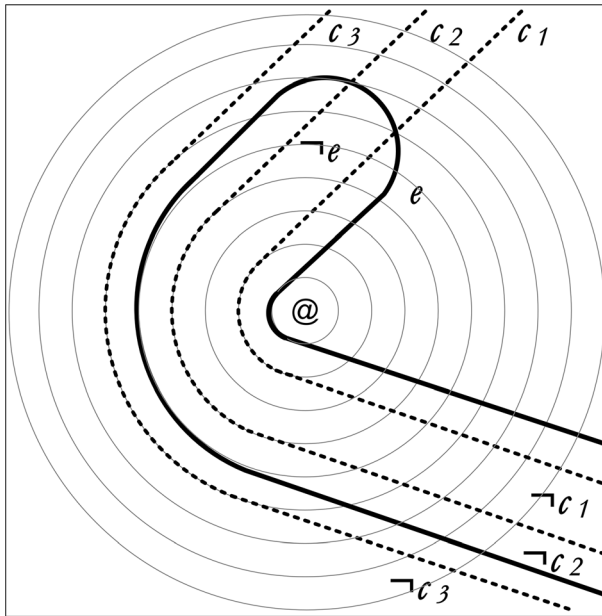
Fig. 1 $\neg O(c_1) \square \mapsto \neg O(e)$ is made true because all $\neg O(c_1)$-worlds closest to @ are also $\neg O(e)$ worlds; $\neg O(c_3) \square \mapsto \neg O(e)$ is false because some $\neg O(c_3)$-worlds closest to @ are not $\neg O(e)$ worlds (left of the third dotted line). But $\neg(O(c_1) \wedge O(c_3)) \square \mapsto \neg O(e)$ is made true

its right-hand side are all worlds in which c_1 occurs, and the worlds to its left-hand side are all worlds in which c_1 does not occur.

The scenario depicted by Fig. 1 renders true the counterfactual $\neg O(c_1) \square \mapsto \neg O(e)$ for the following reason. All closest non-c_1-worlds (all of which are located behind the inner edge of the second circle from @) are worlds in which e does not occur either. The same is the case for the counterfactual $\neg O(c_2) \square \mapsto \neg O(e)$. All closest non-$c_2$-worlds (all of which are located at the inner edge of the fourth circle from @) are worlds in which e does not occur either. However, $\neg O(c_3) \square \mapsto \neg O(e)$ is rendered false by the scenario. All the closest non-c_3-worlds are located behind the third dotted line at the inner edge of the sixth circle from @. All of these closest non-c_3-worlds are e-worlds. In other words, Lewis' second condition "...some $O(\phi)$-world where $O(\psi)$ holds is closer (to w) than is any $O(\phi)$-world where $O(\psi)$ does not hold" is violated.

As a result, whereas c_1 and c_2 are actual causes of e under **Causation** $_{Lewis}$, c_3 is not. Nevertheless, all of the following counterfactuals are true under Lewis' evaluation procedure: $\neg(O(c_1) \wedge O(c_3)) \square \mapsto \neg O(e)$, $\neg(O(c_2) \wedge O(c_3)) \square \mapsto \neg O(e)$, $\neg(O(c_1) \wedge O(c_2) \wedge O(c_3)) \square \mapsto \neg O(e)$. The reason is that, for instance, the closest $\neg(O(c_1) \wedge O(c_3))$-worlds are worlds in which either only $\neg O(c_1)$, or only $\neg O(c_3)$, or both $\neg O(c_1)$ and $\neg O(c_3)$ is/are true. In Fig. 1, obviously the first scenario is the case so that the closest $\neg(O(c_1) \wedge O(c_3))$-worlds are each identical to some closest $\neg O(c_1)$-world. But as just mentioned above, all of these worlds are also $\neg O(e)$-worlds. Hence, $\neg(O(c_1) \wedge O(c_3)) \square \mapsto \neg O(e)$ is true at @.

The question is what these counterfactuals actually express. A natural interpretation is that conjunctions of propositions stating the occurrence of events such as '$O(c_1) \wedge O(c_2) \wedge O(c_3)$' often refer to events themselves; events that are 'summed' from the individual events mentioned by the propositions. As Lewis himself pointed out, "...events are at any rate more amenable to summation than chairs are: a war may be the sum of its battles, a conference may be the sum of its sessions." (Lewis 1986c, 260)[8] In light of the fact that summed events such as battles, conferences, presidential elections, wars, and independence day celebrations are very much in agreement with our everyday and scientific causal talk, Lewis' belief in their existence was certainly justified.[9]

The kinds of events that Lewis refers to in the explication of his theory of causation look a lot like summed events. As he says, his theory is intended to analyze "...causation among *events*, in the everyday sense of the word: flashes, battles, conversations, impacts, strolls, deaths, touchdowns, falls, kisses, and the like." (Lewis 1973a, 558) At least some conversations and impacts are probably scattered, non-contiguous, and summed. Hence, Lewis presumably recognizes at least some summed events as potential causes.

To distinguish summation, which is a relation between events, from conjunction, which is a logical connection between propositions, I introduce the sign '★' as a symbol for the summation function. The claim that an event γ counterfactually depends on the summation of events ϕ and ψ then reads '$\neg O(\phi \star \psi) \square\!\!\rightarrow \neg O(\gamma)$'. The principle of event summation can be expressed as follows:

Event summation: Necessarily, $O(\phi \star \psi)$ if, and only if, $O(\phi) \wedge O(\psi)$.

In other words, necessarily, whenever events ϕ and ψ occur, then also a summed event $\phi \star \psi$ occurs.[10] Given **Event Summation**, $\neg O(\phi \star \psi) \square\!\!\rightarrow \neg O(\gamma)$ is equivalent to $\neg(O(\phi) \wedge O(\psi)) \square\!\!\rightarrow \neg O(\gamma)$.

With these assumptions in the background, the counterfactuals $\neg(O(c_1) \wedge O(c_3)) \square\!\!\rightarrow \neg O(e)$, $\neg(O(c_2) \wedge O(c_3)) \square\!\!\rightarrow \neg O(e)$, $\neg(O(c_1) \wedge O(c_2) \wedge O(c_3)) \square\!\!\rightarrow \neg O(e)$ in conjunction with **Causation** $_{Lewis}$ imply that all of $c_1 \star c_3$, $c_2 \star c_3$, and $c_1 \star c_2 \star c_3$ are causes of e. Does this implication pose a problem for Lewis' theory or not?

On the one hand, one may feel an intuition that it is unproblematic because in Lewis' view the distinction between the 'cause' of an event from, for instance, mere 'causal factors' or 'causal conditions' does not make much sense. Lewis was merely concerned "with the prior question of what it is to be one of the causes

---

[8] At the same time, Lewis did not take a stand on whether event summation works unrestrictedly: "But I leave open the question whether several events, however, miscellaneous, always have another event as their sum." (Lewis 1986c, 260)

[9] It should be noted here that Lewis' policy towards event summation is analogous to claims in his work on set theory and object summations. About a summation of a trout and a turkey he says that "...it is neither fish nor fowl, but it is nothing else: it is part fish and part fowl." (Lewis 1991, 80) For further questions about the (im)possibility of event summation, cf. Sect. 4.

[10] Note that this formulation goes beyond Lewis in specifying *unrestricted* event summation. Restrictions such as relations of 'contiguity' can always be added, of course. For the main argument of this article it suffices to presuppose the existence of at least some sufficiently large summed events.

(unselectively speaking)." (Lewis 1973a, 165) An example often used in the literature is the striking of the match that is the 'cause' of the fire, whilst the oxygen in the air and the absence of rain are mere 'causal factors' or 'causal conditions'. Hence, for Lewis, the resulting abundance of causes may at best be a problem for causal explanation, whilst our natural ontological intuitions may be left unchallenged by a large number of causes for a given effect.

On the other hand, one may also feel an intuitive pull against too comprehensive ontic causes. For one thing, discriminating causes from causal conditions looks like a different challenge than distinguishing causes without irrelevant parts from causes with irrelevant parts. Hence, the abundance Lewis describes in the above quote is a different one than the one encountered in the summed events $c_1 \star c_3$, $c_2 \star c_3$, and $c_1 \star c_2 \star c_3$. Secondly, out of independent reasons, one may be drawn to accepting certain fundamental non-redundancy and simplicity principles as they are applied in various branches of metaphysics and science. If applied to the case of causation, these principles may suggest the non-existence of redundant causes (such as the summed events $c_1 \star c_3$, $c_2 \star c_3$, and $c_1 \star c_2 \star c_3$ relative to e in Fig. 1). In the remainder of this section, I will argue that the latter intuitions eventually supersede the former ones, and redundant causes are indeed problematic in an ontic, and not merely an explanatory, sense.

As a first step towards this conclusion, note that on Lewis' analysis the number of causes for many effects is indeed quite large. For instance, two seemingly inconsistent conclusions are reached with respect to event c_3 in Fig. 1. Whereas above it was diagnosed as not causally related to e due to the falsity of $\neg O(c_3) \square\rightarrow \neg O(e)$, it is now described as a part of the cause $c_1 \star c_3$ of e. As it can be shown, the following supplementation law holds in general.

Supplementation 1: For all actual events ϕ, ψ, γ : if $\neg O(\phi) \square\rightarrow \neg O(\psi)$ and some $\neg O(\phi)$-world is closer to @ than is any $\neg O(\gamma)$-world, then $\neg(O(\phi) \wedge O(\gamma)) \square\rightarrow \neg O(\psi)$.

Proof   The closest world in which $\neg(O(\phi) \wedge O(\gamma))$ is true (from the perspective of @, which is a ψ-, ϕ-, and γ-world) is a world, in which either $\neg O(\phi)$ or $\neg O(\gamma)$ is true (by De Morgan's law). However, if some $\neg O(\phi)$-world is closer to @ than is any $\neg O(\gamma)$-world, then it will be this $\neg O(\phi)$-world that is the closest satisfying the proposition $\neg(O(\phi) \wedge O(\gamma))$. Hence, whether or not $\neg O(\psi)$ counterfactually depends on $\neg O(\gamma)$, the counterfactual $\neg(O(\phi) \wedge O(\gamma)) \square\rightarrow \neg O(\psi)$ will always be true, given that $\neg O(\phi) \square\rightarrow \neg O(\psi)$ is true. $\square$

The following actual example illustrates what is going on here: Suppose that, had the bombing of Guernica (call this actual event 'c') not occurred, the destruction of Guernica (call this event 'e') would not have occurred either. Then most likely it will also be true that, had the bombing of Guernica *and* the simultaneous implosion of a distant planet (call this actual event 'c^*') not occurred, the destruction of

Guernica would not have occurred either.[11] The reason is that, according to Lewis' similarity criteria for the ordering of possible worlds, some non-bombing (or $\neg c$-) world is almost certainly closer to the actual world than is any non-implosion (or $\neg c^*$-) world. (Lewis 1979a, 472) adopts the following hierarchy of similarity measures between possible worlds for the evaluation of counterfactuals.

1. It is of first importance that worlds minimize big, widespread violations of actual laws of nature (that minimize big miracles, in other words).
2. It is of second importance that worlds maximize spatio-temporal regions that are identical with our world in matters of particular facts.
3. It is of third importance that worlds minimize small, punctual violations of actual laws of nature as well (small miracles).
4. It is of fourth importance that worlds maximize approximate similarity in matters of particular facts with our world.

As c has a tiny spatio-temporal extension in comparison to the enormously large c^* that occurred also, then if the same laws hold at c and c^*, it will take much more adjustment of particular fact, or a much larger "miracle", in order to prevent the occurrence of c^* than it takes to prevent the occurrence of c relative to the actual world. In other words, the closest non-c^*-world and the closest non-c-world are not at all on a par with respect to a match of particular fact with @. Therefore, according to criterion (2), any closest non-c^*-world (that otherwise matches @) is farther away than at least some closest non-c-world (that otherwise matches @). As a consequence, both the counterfactuals $O(c) \wedge O(c^*) \square\!\!\rightarrow O(e)$ and $\neg(O(c) \wedge O(c^*)) \square\!\!\rightarrow \neg O(e)$ are true at @.

Under this assumption, the referent of the expression "the bombing of Guernica *and* the simultaneous implosion of a distant planet", i. e. of $O(c) \wedge O(c^*)$, is considered a summed but genuine event, namely $c \star c^*$. As a consequence, the left-hand side of **Causation** $_{Lewis-R}$ is satisfied by the summed event $c \star c^*$ with respect to e, which is just to say that $c \star c^*$ proves to be a cause of e.

Since what is happening with a distant planet has at least no immediate impact on what happens on earth, the complex event consisting of the bombing of Guernica *and* the simultaneous implosion of a distant planet contains irrelevant stuff. But, again, is it actually a problem for the counterfactual theory if it identifies causes containing irrelevant detail?

One reason for believing that it is indeed a problem is our everyday practice of assigning causal relationships to pairs (or ordered sets) of events. The latter is intuitively measured by following the trajectories of events in space-time. If we ask ourselves what caused the golden ball to fall into the well, we typically trace back

---

[11] Through some odd causal chain the implosion might have led to the destruction of Guernica eventually, of course. However, in counterfactual theories of causation it is usually presupposed that the actual events in question cannot occur at all times and still be the same event. For instance, the destruction of Guernica as it occurred on April 26, 1937 could perhaps have occured also on April 27, 1937. But it could not have occurred on April 26, 2037. Consequently, the odd causal chain that leads to the destruction of Guernica at some remote point in the future is not one that has e as its effect, i.e. the destruction of Guernia on April 26, 1937.

the ball's trajectory to the princess' hand movement. A trajectory leading to the hand movement *and* a battle raging between the king's army and an enemy on a distant land is far less intuitive. This may already indicate an ontic, and not merely an explanatory, distinction.

A second and stronger reason for believing that causal redundancy is a problem is that it violates one of the most basic and widely accepted constraints on ontology and science. The constraint is often dubbed "Ockham's razor" alluding to the 14th century scholar William of Ockham who famously demanded: *Numquam ponenda est pluralitas sine necessitate* ["Plurality must never be posited without necessity"]. The fundamental idea has received various formulations in many different philosophical and scientific contexts (cf. Baker 2016). A version pertaining to causation is found in Newtons statement that "[w]e are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances." (cited in Hawking 2003, 171). Whilst Lewis does not invoke simplicity in the service of reducing the number of causes, he has embraced the general idea in his conception of a law of nature. In his view, a law of nature is a contingent regularity in a world w that is a theorem or an axiom in any of the best deductive systems true at w (Lewis 1973b, 73; Lewis 1994, 478). A best deductive system is one that balances best parsimony and strength.

Parsimony principles have sometimes been justified on *a priori* grounds. One such defense claims that redundancy violates our aesthetic sense and our standards of rationality. Naturalistic justifications have cited as empirical evidence the patterns of acceptance and rejection of competing theories by working scientists (cf. Baker 2016, sections 3-4.) Given that these justifications are cogent, it would be peculiar if parsimony constrained our ontological and scientific attitudes, but turned out to be inapplicable to our theories of causation. Or, in other words, if the justifications of the parsimony principle are convincing in other fields, they ought to guide us in our attempt to distinguish causes from non-causes as well. One could formulate this idea in an Occhamian spirit as *Non sunt multiplicanda causas sine necessitate*, or "Causes are not to be multiplied without necessity".[12] If we take this principle seriously, the summed events $c_1 \star c_3$, $c_2 \star c_3$, and $c_1 \star c_2 \star c_3$ should not turn out to be causes relative to e in the situation illustrated by Fig. 1.

Even if this point about redundancy is granted, a fictional critic might still object that **Supplementation 1** is true but irrelevant for causal analysis. After all, Lewis' himself did not make explicit how the negative counterfactual within a causal analysis would have to look like when the antecedent is a conjunction. The critic may claim that Lewis would have argued that, when a summed event $c_1 \star c_2$ is tested for causation, the relevant counterfactual has to take on the following form instead: $\neg O(c_1) \wedge \neg O(c_2) \square \rightarrow \neg O(e)$ (in contrast to $\neg(O(c) \wedge O(c^*)) \square \rightarrow \neg O(e)$). If altered accordingly, **Supplementation 1** is demonstrably false.

Unfortunately though, the fictional critic's proposal does not save Lewis' theory in general. The reasons is that, if the negative counterfactual is demanded to be formed in this way, a related problem occurs for Lewis' theory of causation.

---

[12] For a formal definition of this principle, see Sect. 5.
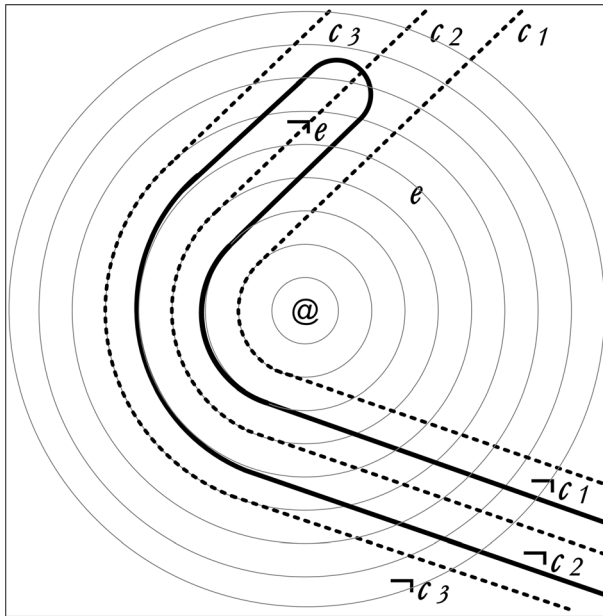
Fig. 2 $\neg O(c_2)\square\!\!\rightarrow \neg O(e)$ is made true because all $\neg O(c_2)$-worlds closest to @ are also $\neg O(e)$ worlds; $\neg O(c_1)\square\!\!\rightarrow \neg O(e)$ is false because some $\neg O(c_1)$-worlds closest to @ are not $\neg O(e)$ worlds (left of the first dotted line). But $\neg(O(c_2) \wedge O(c_1))\square\!\!\rightarrow \neg O(e)$ is made true

Consider the case illustrated by Fig. 2, in which $\neg O(c_2)\square\!\!\rightarrow \neg O(e)$ is true, $\neg O(c_1)\square\!\!\rightarrow \neg O(e)$ is false, but $\neg O(c_1) \wedge \neg O(c_2)\square\!\!\rightarrow \neg O(e)$ is true. Again, although c_1 is obviously not causally related to e, the summed event $c_1\star c_2$ qualifies as a cause of e. However, given that c_1 is completely irrelevant for e, $c_1\star c_2$ should not be characterized as a cause. The corresponding supplementation law is the following:

Supplementation 2: For all actual events ϕ, ψ, γ : if $\neg O(\phi)\square\!\!\rightarrow \neg O(\psi)$ and all $\neg O(\phi)$ worlds closest to @ are also $\neg O(\gamma)$ worlds, then $\neg O(\phi) \wedge \neg O(\gamma)\square\!\!\rightarrow \neg O(\psi)$.

We take **Supplementation 1**, or **Supplementation 2** respectively, to show that **Causation** $_{Lewis-R}$, and hence **Causation** $_{Lewis}$, is false even for ordinary cases. Counterfactual dependence under the classical definition is not sufficient for causation, because effects typically counterfactually depend on various complex events that contain almost nothing that is causally relevant to the effects in question. If these would be accepted as causes, virtually every event would be part of a cause of almost every other event, which violates some fundamental constraints on our ontological beliefs. Hence, if the complex events referred to by **Supplementation 1** and **Supplementation 2** exist, then true counterfactuals must satisfy at least one additional constraint in order to be causally interpretable. Before this constraint is formulated, the next section considers two potential objections against the relevance

of **Supplementation 1** and **Supplementation 2** for counterfactual theories of causation.[13]

4 Event ontology

As mentioned in the previous section, **Supplementation 1** and **Supplementation 2** are based on the assumption shared by Lewis that summed events exist. However, as a consequence, Lewis' approach postulates too many causes.

A potential reaction to this diagnosis by a fictional opponent could be to recognize that Lewis' theory issues weird causal dependencies on the basis of summed complex events, and at the same time to point out that the 'weirdness' here stems from a very unnatural event ontology involving outlandishly summed events that clearly transcend Lewis' tolerance zone. Events that are radically scattered and non-continuous are simply no candidates for causal claims. Once this weird event ontology is rejected, the problem of large causes disappears.

This reasoning is problematic in at least three respects. First, it is far from clear that summed but scattered events cannot enter into causal relations also in Lewis' sense. It seems perfectly fine to say that the joint attack of the armies of Russia, Prussia, Austria and Sweden caused Napoleon's defeat at Leipzig in 1813. However, the joint attack involved four major non-adjacent battlefields, rendering the causally efficacious battle of Leipzig a summed but scattered event (again, cf. Lewis 1986c, 260).

Secondly, it is doubtful that even acceptable summed colloquial events can escape the problem of large causes. Or in other words, the problem of large causes generally concerns events which intuitively contain irrelevant parts. This tends to go along with being spatio-temporally scattered, as in the above examples, but it is not the same property. The problem also occurs with respect to many non-scattered events. Take as an example the destruction of the Moracae Ficus specimen in the economic botanic collection of Bristol's city museum on 24 November 1940 that presumably non-trivially counterfactually depends on WWII (at least on that part of WWII which does not overlap with, and happened before, the destruction of the specimen). WWII constructed as an extended event could be considered summed from the Invasion of Poland, the Invasion of Belgium, the Invasion of the Netherlands, ... etc. However, the size of this event commonly referred to by 'WWII' is radically at odds with the magnitude of a destruction of a tiny container. Hence, it seems misleading to contend that WWII caused the destruction of the

---

[13] Note that the problem of large causes is implicitly alluded to already by (Schaffer 2003b, 28) when he argues that two parts of a big rock can both be causes of a window shattering even if they are not all "relevant". In Schaffer's view, the reason is that the overdetermining parts are "lawfully yoked". Schaffer seems to have overlooked the gravity and scope of this problem in light of the supplementation laws **Supplementation 1** and **Supplementation 2**, which do not only refer to "lawfully yoked" events. Secondly, the problem has been mentioned explicitly by Sartorio (2006, 378-79). However, her analysis does not go beyond a non-formal version of **Supplementation 1**.

specimen.[14] More likely it was the fire that broke out after a bomb had been dropped on the museum by a German Arado Ar 234 bomber. Perhaps the entire air raid consisting of this Arado's and several other bombers' flight could still count as an intuitive cause of the destruction of the specimen. But WWII cannot. It is simply too large, too extended, and filled with too much irrelevant stuff to be considered the, or a, cause of the specimen's destruction. Nevertheless, it is a non-scattered event, and under the *Lewis Analysis*, the counterfactual ¬*WWII* □→ ¬ *the destruction of the Moracae Ficus specimen* comes out true. Consequently, WWII is described as a cause of the specimen's destruction under **Causation** $_{Lewis}$.

The example can be radicalized at will. If the claim that 'WWII caused the destruction of the specimen' still appears intuitive to some extent, one can broaden the claim, for instance, into: 'The entire history of the universe from the year 1900 until 24 November 1940 caused the destruction of the specimen'. Again, under the *Lewis Analysis* the corresponding counterfactual ¬ *Complete universe 1900 until 24 November 1940* □→ ¬ *the destruction of the Moracae Ficus specimen* comes out true. But the radicalized claim 'The entire history of the universe from the year 1900 until 24 November 1940 caused the destruction of the specimen' simply seems false. The reasons is that just about everything of what happened 'within' the universe during this period had nothing to do with the destruction of the specimen. Only an extremely tiny and almost negligible part of this huge event had some causal connection to the specimen. But again, note that the entire history of the universe from the year 1900 until 24 November 1940 is a non-scattered event because it is entirely continuous. Hence, the problem of large causes cannot simply be a matter of an event ontology that permits scattered and non-continuous events.

Thirdly, the problem of large causes can be constructed for certain non-scattered non-summed events as well. Suppose that the 1997 Superbowl game was such a non-summed event. Then the relocation of some carbon atom in the turf of San Diego's Qualcomm Stadium counterfactually depends on the 1997 Superbowl game. However, it seems wrong to describe the game as having caused the relocation. More plausibly, it was quarterback John Elway's run over the turf (assuming that he out of all players was the only one who actually directly affected the atom's relocation). Or simply consider the complete state of the universe preceding the moment of the atom's relocation. The entire state as a complete event is continuous and perhaps non-summed, but it surely is much too large to be considered the cause of the relocation of the atom. The problem here stems from an almost total lack of "fit" between cause and effect. Moreover, the general version of the universe example implies that almost anything is part of a cause of anything else.

In light of these examples, it becomes clear that the "weirdness" of the counterfactual dependencies that constitute the problem of large causes does not

---

[14] Note that it does not help at this point to simply point out that causes must always precede their effects. For a Humean, the arrow of time is not a primitive relation and cannot be used as a criterion for causal analysis. Moreover, there may be reasons why a Humean would want to at least potentially allow causes that occur later than their effects.

stem from an unnatural event ontology involving outlandishly summed events. Rather, it stems from the *Lewis Analysis* for counterfactuals.

Note that these cases of WWII and the Superbowl game cannot simply be accommodated by the distinction of causation and causal explanation. The reason why WWII and the Superbowl game are not good candidates for the causes of the specimen destruction or the relocation of the atom is quite different from the one that leads us to hesitate in calling the birth of the driver's paternal grandmother a cause of his fatal accident (cf. Lewis 1986a, 216). Our reluctance to cite the birth as a cause is grounded in various pragmatic constraints that we tend to apply in the context of causal explanations. However, we can still force appreciation of the causal relevance of the birth by consciously canceling these ordinary pragmatic factors. In contrast, canceling the same pragmatic constraints from our explanation in the cases of WWII and the Superbowl does not remove our reluctance to call them causes of the specimen destruction or the relocation of the atom. And it is here where the problem lies.

A second objection may try to argue that Lewis does not face the problem of large causes as he identifies events by their essential and intrinsic properties that are had by space-time regions. As he says: "We may classify events by their essences, stating conditions that a region must satisfy if that event is to occur there." (Lewis 1986c, 247) Moreover, "[a]n event is a localized matter of contingent fact. ...[It] occurs in a particular spatiotemporal region." (*op. cit.*, 243)[15] Even if summed events occupy space-time regions, it could perhaps be argued that summations of intrinsic events do not create events with essential and intrinsic properties as well. Hence, there are no summed events in Lewis' sense.

It seems difficult to prove rigorously that summation is closed under the class of events with essential and intrinsic properties. On the other hand, it also seems hard to show that the summed event of Suzy's throw of a stone and Bill's throw of a stone is not identifiable by its essential and intrinsic properties. How are we to resolve this issue? Lewis' own answer is clear. Even though he left open whether event summation is unrestricted, he accepted at least some event summations as genuine events (Lewis 1986c, 260). Hence, at least in Lewis' view, the fact that events are identified by their essential and intrinsic properties does not categorically exclude the existence of summed events. And this is all what's required for the problem of large causes to raise its head.

A third objection to the conclusion based on **Supplementation 1** and **Supplementation 2** may point out that the problem of large causes is too general

---

[15] It should be noted here that Lewis was not always consistent about the logical structure of an event. In some places, he defined events as *instantiations of properties by space-time regions* (cf. Lewis 1986c, 243). Call this the "occurrant" notion of an event. Events understood in this way are neither individuals nor classes (=properties); rather they are those things that correspond to sentences in first-order logic. In other places, Lewis has characterized events as *properties, which are classes of space-time regions* (cf. Lewis 1986c, 245). Call this the "property" or "class" notion of an event. Events understood in this way are neither individuals nor correspondants of sentences in first-order logic; rather, they are those structures corresponding to predicates in first-order logic. Without being able to discuss this ambiguity in Lewis' writings further, it should be clear that in this paper I presuppose that the "occurrant" notion of an event is at the foundation of Lewis' counterfactual theory of causation.

to be a problem specifically for counterfactual theories of causation. The point is that all causal analyses must be relativized to an appropriate modeling apparatus without it always being clear what "appropriate" means in this context. Type-level analyses must be relativized to a set of variables that are appropriate for causal modeling, and token-level analyses must be relativized to an appropriate event ontology or an appropriate assignment of values to variables.[16] But everybody's problem cannot be brought up against a specific theory of causation.

First, it should be noted that it is unclear that this is the same problem as the problem of large causes. It is true, of course, that the latter problem is one aspect of the problem of selecting appropriate variables and values. But it is not obvious at all that the two problems are the same. Hence, the problem of large causes may still be significant on its own right.

Secondly, this potential objection ignores an important distinction between reductionist Humean and non-reductionist theories of causation. As an example, interventionist theories do not aspire to provide a reductive analysis of causation. Theories of this kind are primarily concerned with the justification of causal reasoning. As a consequence, a continuous re-adjustment of the values and variables in the sense of determining an "appropriate modeling apparatus" is admissible within the theory. Prior causal knowledge can well play a substantial part in these re-adjustment processes. The same holds for at least some probabilistic theories of causation, for which a continuous re-adjustment of the pertinent variables is an admissible procedure.

Reductionist counterfactual theories, in contrast, cannot allow the event ontology to be fixed based on prior causal assumptions, i.e. knowledge allowing to distinguish between admissible and non-admissible events within a causal modeling apparatus. Their aim is to produce adequate causal analyses in terms of counterfactual conditionals. In this sense, the "appropriate modeling apparatus" presupposed by counterfactual theories can only be constrained by general non-causal criteria. Examples of such criteria would be: "No supernatural events are to be admitted" or "Only events with absolute times are to be admitted" etc. However, as the discussed examples made clear, criteria of this kind do not by themselves rule out the problem of large causes.

A simple way to avoid the problem of large causes would, of course, be to embrace radical atomism and rule that there are no such things as large events. Lewis' own position might perhaps be interpreted in this way, as it is summarized in his famous quote:[17]

> Humean supervenience (...) is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. (Lewis 1986d, ix)

---

[16] For instance, Woodward (2016) has explicitly discussed a related problem which he calls the "problem of variable choice".

[17] But note that radical atomism may be at odds with the passage from (Lewis 1973a, 558) already quoted in Sect. 2 "I shall confine myself to causation among *events*, in the everyday sense of the word: flashes, battles, conversations, impacts, strolls, deaths, touchdowns, falls, kisses, and the like."

Such a radical atomism combined with the rejection of event summations would obviously avoid the problem of large causes. But it would also avoid any causation involving flashes, battles, conversations, impacts, strolls, deaths, touchdowns, falls, kisses, and the like. Even if flashes, battles and so on can be reduced to configurations of atomic events, no kiss is identical to an event with the extension of a single atomic space-time point. This forces counterfactual theories into a dilemma. Either only fundamental atomic events are allowed into the ontology. Then the problem of large causes is avoided, but also almost all our causal talk in colloquial as well as scientific contexts is misguided.[18] Or summed or macro-events are admitted into the ontology. Then the problem of large causes will almost certainly occur at some point if the *Lewis Analysis* is presupposed.

Since it seems a bit radical to recognize only fundamental atomic events as potential causes and effects, the only solution for counterfactual theories is to try to find some ways around the the problem of large causes using their own original toolbox. The next section indicates how such a strategy could look like.

5 Antecedent minimization

A potential solution to the problem of overly large causes discussed in the previous sections is to introduce a constraint ensuring that counterfactuals are causally interpretable only if they have a *minimized* antecedent. An improved counterfactual theory of causation between actual events could then take the following form:

Causation $_{AntMin}$: ϕ is a cause of ψ if, and only if, $O(\phi)$ is counterfactually minimized with respect to $O(\psi)$.

The problem is, of course, to make the notion of "counterfactual minimization" formally precise. If any event ϕ is either atomic or a summed event $\phi_1 \star \ldots \star \phi_n$, i. e. if ϕ is either atomic or coextensive with some summation of events $\phi_1, \ldots, \phi_n$, the definition is fairly straightforward. Say that a summing part ϕ_k of an event ϕ is any summation with $n > 0$ summed events that is attained by removing one or more events from the event summation that is ϕ. Then counterfactual minimization can be described as follows.[19]

Counterfactual Minimization 1: $O(\phi)$ is counterfactually minimized with respect to $O(\psi)$ if, and only if, for all ϕ_k, if ϕ_k is a summing part of ϕ, then $\neg O(\phi_k)\square\!\!\rightarrow \neg O(\psi)$.

Note that, if an event ϕ has no summing parts, satisfaction of **Causation** $_{Lewis}$ with respect to another event ψ is sufficient for ϕ to cause ψ. To see how **Causation**

---

[18] It should be pointed out that nihilism about composition has similar consequences. See, for instance, (Merricks 2001, ch. 3).

[19] As stated, **Counterfactual Minimization 1** presupposes that summation parthood is a reflexive relation. If it were not reflexive, **Causation** $_{AntMin}$ would have to be altered as follows: "ϕ is a cause of ψ iff (i) $\neg O(\phi)\square\!\!\rightarrow \neg O(\psi)$ and (ii) $O(\phi)$ is counterfactually minimized with respect to $O(\psi)$."

AntMin works in detail, suppose that the bombing of Guernica consisted (as it actually did) of the raid of the Aviazione Legionaria (call it 'c^{**}') and the raid of the Condor Legion (call it 'c^{***}'). Suppose further that, if one of the two raids would not have occurred and the other would still have, Guernica would not have been destroyed but turned into rubble "merely" by half. Then both $\neg O(c^{**}) \square\rightarrow \neg O(e)$ and $\neg O(c^{***}) \square\rightarrow \neg O(e)$ are true, and c ($= c^{**} \bigstar c^{***}$) is characterized by **Causation** *AntMin* as a cause of e. In contrast, the complex event $c \bigstar c^*$, the summation of the bombing of Guernica and the implosion of a distant planet, is no longer a cause of e because $\neg O(c^*) \square\rightarrow \neg O(e)$ is false and **Causation** *AntMin* is violated. This is all as desired.

Unfortunately, however, **Counterfactual Minimization 1** has a serious defect.[20] To see why, note that saying that the bombing of Guernica consisted of the raids of the Aviazione Legionaria and the Condor Legion is a gross simplification. More precisely, the bombing consisted of a huge number of local events including, for instance, the explosion of a bomb with serial number 16'429 released by a plane of the Condor Legion (call this event 'c^{****}'). However, the claim that $\neg O(c^{****}) \square\rightarrow \neg O(e)$ is clearly false. As a consequence, c no longer qualifies as a cause of e under **Causation** *AntMin*. This is a serious problem for our analysis, because c was in fact the cause of e and should have qualified as such under an adequate theory of causation.

This Sorites Paradox kind of problem can only be countered if counterfactual minimization is defined by a condition like **Counterfactual Minimization 2**.

Counterfactual Minimization 2: $O(\phi)$ is counterfactually minimized with respect to $O(\psi)$ if, and only if, for all ϕ_k, if ϕ_k is an *essential* summing part of ϕ, then $\neg O(\phi_k) \square\rightarrow \neg O(\psi)$.

An essential summing part of an event ϕ would be a part whose removal would affect the identity of ϕ. The idea of **Counterfactual Minimization 2** in combination with **Causation** *AntMin* is then that an event ϕ is a cause of an effect ψ if, and only if, all of the parts of ϕ that jointly constitute ϕ's identity are counterfactually relevant for ψ. Of course, within a Humean metaphysics it is always possible that the very same event has very different effects in different circumstances. For instance, if the bombing of Guernica would have happened in 2020, the attacking country would have immediately faced a massive response by NATO. Hence, if an event ϕ has ϕ' and ϕ'' as essential parts and ψ as its effect in the sense of **Counterfactual Minimization 2** and **Causation** *AntMin*, it does not follow that ϕ has effect ψ essentially. The criterion for event identity does not line up with the event's causal relationships in this way.

However, in a particular world or situation **Counterfactual Minimization 2** and **Causation** *AntMin* do align the identity condition of an event with its causal relations to other events. The reason is that, as suggested in Sect. 3, an event is the cause of

---

[20] The problem discussed here is complementary to the problem of a large event e lacking counterfactual dependence on a large event c, even though c and e are divisible into parts in such a way that every part of e counterfactually depends on some part of c (cf. Lewis 1986e, 172-175). Lewis' notion of a "piecemeal causation" may solve this problem and describe c as a cause of e, but it does not offer a solution to the problem of large causes and essential parts of causes.

an effect only if it is not redundant relative to this effect in a particular world. The challenge for a counterfactual theory of causation then consists in providing criteria for determining the essential summing parts of an event without recurring to prior causal knowledge.

The challenge of offering criteria for the essential summing parts of an event is connected to the problem of determining which events are ontologically fundamental. In particular, note that any event that has no summing parts is counterfactually minimized with respect to any other event. This implies that **Causation** $_{AntMin}$ as an adequate condition for actual causation depends on the graininess of the atomic events presupposed by the analysis, i. e. events that have no summing parts. For example, suppose there is an event u_e that is the complete state of the universe preceding e. Suppose further that the fundamental ontology of events used for a causal analysis based on **Causation** $_{AntMin}$ does not contain any conjunctive part of u_e, so that u_e is not summed from any event. Then u_e will satisfy **Causation** $_{AntMin}$ with respect to e. Intuitively however, almost everything "within" u_e is completely causally irrelevant for e.

One way to counter this problem would be to adjust the fundamental event ontology presupposed until the causal analysis fits our intuitions about causal relevance. Under such a solution, the question about the essential parts of an event might lose its relevance. But in the same way as before, it would not be admissible to include prior causal information into the criteria guiding the adjustment. Another strategy would perhaps be to introduce the notion of an unrestricted division of events into the definition. According to this idea, any event is divisible into any number of subevents each occupying a different part of the spatio-temporal extension of the event. If the symbol '$\subset$' is used to refer to the relation subevents bear to their subsuming events, counterfactual minimization can be made precise as follows.

Counterfactual Minimization 3: $O(\phi)$ is counterfactually minimized with respect to $O(\psi)$ if, and only if, for all ϕ', if $\phi' \subset \phi$ then $\neg O(\phi') \square\!\!\rightarrow \neg O(\psi)$.

This proposal seems to work with respect to the bombing of Guernica, given it is divided into the two raids, and with respect to the complex event summed from the bombing and the implosion of a distant planet since only the former satisfies **Counterfactual Minimization 3** with respect to the destruction of Guernica. But it is easily seen that, ultimately, this proposal is of little help as well. In particular, suppose that division is infinite and that there are no atomic events at which division naturally stops.[21] In this case, the truth value of the right-hand side of **Counterfactual Minimization 3** cannot be determined by any finite number of analytic steps, and the causal relation between any two events becomes indeterminable under **Causation** $_{AntMin}$.

On the other hand, if there are atomic events and event division has a natural lower boundary (i.e. it is not infinite), the Sorites Paradox kind of problem

---

[21] For an argument stating that there is no evidence for fundamentality, see Schaffer (2003a).

diagnosed for **Counterfactual Minimization 1** is likely to reappear again. For many summed events ϕ, it will be the case that their effects do not counterfactually depend on any subevents of ϕ. Just as before, the solution can only be to bind the notion of counterfactual minimization to the notion of an essential subevent.

Counterfactual Minimization 4: $O(\phi)$ is counterfactually minimized with respect to $O(\psi)$ if, and only if, for all ϕ', if $\phi' \subset_{essential} \phi$ then $\neg O(\phi') \square\!\!\rightarrow \neg O(\psi)$

If this path is chosen, the challenge for making sense of **Counterfactual Minimization 4** will be to give a satisfactory definition of the notion of an essential subevent. If the definition should be of any help to **Causation** $_{AntMin}$, it has to be spelt out in non-causal terms.

One may speculate that this task is solvable on the basis of Lewis' similarity criteria for possible worlds (cf. Sect. 3). In particular, perhaps the following law is implied by the criteria: For any actual event ϕ, any essential subevent ψ of ϕ and any non-essential subevent γ of ϕ, the closest $\neg\gamma$-worlds are necessarily closer to @ than any closest $\neg\psi$-world. Unfortunately, there are at least two problems with this law in the context of **Counterfactual Minimization 4**. First, its implication by Lewis' criteria is not obvious. Its truth presupposes that the absence of any essential subevent of any actual event necessarily violates more laws of @ than does the absence of any non-essential subevent. Why should this be so? The answer seems far from clear.

Moreover, even if the law were true, it still would not provide an independent criterion for distinguishing the essential from the non-essential parts of an actual event. For how are we to draw the line between those absence worlds that are "sufficiently far away" for them to be lacking essential subevents of the initial event from those that are "still close enough" to be lacking only inessential subevents? No criterion seems to be forthcoming that would not itself draw on prior causal knowledge. Consequently, with Lewis' similarity criteria alone no satisfactory definition of the notion of an essential subevent is achieved that would help to rigorously specify **Causation** $_{AntMin}$.[22]

With the conclusions just stated, we arrive at the most important implication of the problem of large causes. If the counterfactual theory of causation is to specify at least sufficient conditions for causation in ordinary cases, it must offer an analysis of what it is for something to be an essential summing part and/or an essential subevent. And it must do so without drawing on prior causal knowledge. If no such analysis is forthcoming, the counterfactual theory of causation is doomed. It fails to

---

[22] Note also that Lewis' general claim that events can be classified by their essences (Lewis 1986c, 247; cf. also Sec. 4) does not by itself solve the question what is an essential subevent of an actual event. Lewis simply does not offer an explicit account of an essential subevent.

offer a correct analysis of many ordinary cases unless it substantially departs from Lewis' original intention, which was to provide a reductive account of causation.[23]

Now suppose for the sake of the argument that **Counterfactual Minimization 4** can eventually be amended by an adequate analysis of an essential summing part and/or an essential subevent. Then a further objection may emphasize that it is still too restrictive to count as an adequate theory of causation. In much of our ordinary talk, we tolerate irrelevant portions of causes, and we usually do not distinguish between non-redundant and slightly redundant causes. For instance, perhaps we are willing to accept that the complex event of the bombing of Guernica and the howling of the Guernica sirens caused the destruction of Guernica. That is perhaps because we are not completely sure whether the loudness of the sirens might have shattered a few windows even shortly before the bombing, or whether the sirens actually helped the Condor Legion to locate its target, or whether the role of the sirens in the event will ever be fully understood. In consequence, **Counterfactual Minimization 4** renders at large portion of our ordinary talk literally false, which calls its adequacy into question.

It seems to me that the objection has a point with respect to causal explanation. A theory of explanation that disqualifies almost all actual explanatory practice requires fixing. It is a bit less clear that the objection also challenges **Counterfactual Minimization 4** if interpreted as a theory of causation. The mere fact that our ordinary talk is often flexible and loaded with shifting implicit assumptions (cf. Lewis 1979b) does not immediately imply that we should accept something similar for our ontological beliefs. However, if the latter is assumed, it is clear that **Counterfactual Minimization 4** needs to be modified in some way. It will have to allow at least some slightly redundant causes. It should be clear, though, that such a modification would not necessarily hurt the spirit of **Counterfactual Minimization 4** and its main achievement relative to Lewis' original theory.

But still, the overall problem remains. It is unclear how the notion of an essential summing part and/or an essential subevent can be attained within Lewis' theory without re-introducing causal vocabulary at some point. As a consequence, the problem of large causes is a real threat for Lewis' counterfactual theory of causation even for the 'ordinary' cases, for which so far it has widely been considered adequate.

6 Conclusion

After re-constructing the main ideas of Lewis' counterfactual theory of causation (cf. **Causation** $_{Lewis}$), Sect. 3 first spelt out the problem of large causes. The problem

---

[23] As a side remark, note that modern regularity theories of causation face a corresponding problem that as well seems to have not been duly addressed in the literature so far. A minimization constraint is, of course, at the heart of modern regularity theories of causation (also known as "INUS"-conditions; cf. Mackie 1974, 62; Graßhoff and May 2001; Baumgartner 2008, 2009). However, the graininess of the types presupposed in the analysis heavily influences the established causal judgments, which may render the latter intuitively inadequate in some cases.

made clear that **Causation** $_{Lewis}$ is inadequate even from the original counterfactual theorist's perspective. It was argued in Sect. 4 that the problem does not merely stem from an outlandish event ontology, but from the analysis that Lewis offers for counterfactual conditionals. Section 5 then offered a tentative solution that would allow the counterfactual theory of causation to avoid the problem of large causes. However, this solution required a definition of the notion of an essential summing part and/or an essential subevent. Some doubts were raised whether the definition can be attained within Lewis' theory without re-introducing causal vocabulary at some point. As a consequence, the problem of large causes was characterized as a real threat for Lewis' counterfactual theory of causation even for the 'ordinary' cases, for which it is still widely considered adequate. In other words, the counterfactual theory may eventually lack the resources to offer an adequate analysis of causation without substantially departing from Lewis' original intention, which was to provide a reductive account of causation.

References

Albert, D. (2012). Physics and chance. In Y. Ben-Menahem & M. Hemmo (Eds.), *Probability in physics* (pp. 17–40)., The frontiers collection, chapter 2 Berlin: Springer.

Baker, A. (2016). Simplicity. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2016 ed.). Metaphysics Research Lab, Stanford University.

Baumgartner, M. (2008). Regularity theories reassessed. *Philosophia*, *36*(3), 327–354.

Baumgartner, M. (2009). Uncovering deterministic causal structures: A Boolean approach. *Synthese*, *170*(1), 71–96.

Bennett, K. (2003). Why the exclusion problem seems intractable, and how, just maybe, to tract it. *Noûs*, *37*(3), 471–497.

Dowe, P. (1992). Wesley Salmon's process theory of causality and the conserved quantity theory. *Philosophy of Science*, *59*(2), 195–216.

Elga, A. (2001). Statistical mechanics and the asymmetry of counterfactual dependence. *Philosophy of Science*, *68*(S3), S313–S324. https://doi.org/10.1086/392918.

Frisch, M. (2005). *Inconsistency, asymmetry, and non-locality: A philosophical investigation of classical electrodynamics*. Oxford: Oxford University Press.

Frisch, M. (2011). From arbuthnot to boltzmann: The past hypothesis, the best system, and the special sciences. *Philosophy of Science*, *78*(5), 1001–1011.

Graßhoff, G., & May, M. (2001). Causal regularities. In W. Spohn, M. Ledwig, & M. Esfeld (Eds.), *Current issues in causation* (pp. 85–114). Paderborn: Mentis.

Hall, N. (2007). Structural equations and causation. *Philosophical Studies*, *132*(1), 109–136.

Hawking, S. (2003). *On the shoulders of giants*. Philadelphia: Running Press.

Horwich, P. (1993). Lewis's programme. In E. Sosa & M. Tooley (Eds.), *Causation* (pp. 208–216). Oxford: Oxford University Press.

Hume, D. (2000/1748). *An enquiry concerning human understanding: a critical edition*. Oxford: Oxford University Press.

Hüttemann, A. (2013). *Ursachen*. Berlin: De Gruyter.

Kroedel, T. (2008). Mental causation as multiple causation. *Philosophical Studies*, *139*(1), 125–143.

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

Lewis, D. (1973b). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, D. (1973c). Counterfactuals and comparative possibility. *Journal of Philosophical Logic*, *2*(4), 418–446.

Lewis, D. (1979a). How to define theoretical terms. *The Journal of Philosophy*, *117*(13), 427–446.

Lewis, D. (1979b). Scorekeeping in a language game. *Journal of Philosophical Logic*, *8*(1), 339–359.

Lewis, D. (1986a). Causal explanation. See Lewis (1986d), pp. 214–240.

Lewis, D. (1986b). Counterfactual dependence and time's arrow. See Lewis (1986d), pp. 32–52.

Lewis, D. (1986c). Events. See Lewis (1986d), pp. 241–269.

Lewis, D. (1986d). *Philosophical papers* (Vol. II). Oxford: Oxford University Press.

Lewis, D. (1986e). Postscripts to 'Causation'. See Lewis (1986d), pp. 172–213.

Lewis, D. (1991). *Parts of classes*. Hoboken: B. Blackwell.

Lewis, D. (1994). Humean supervenience debugged. *Mind*, *103*(412), 473–490.

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

Loewer, B. (1996). Humean supervenience and laws of nature. *Philosophical Topics*, *24*(1), 101–127.

Loewer, B. (2007). Mental causation, or something near enough. In B. McLaughlin & J. Cohen (Eds.), *Contemporary debates in philosophy of mind* (pp. 243–264). Hoboken: Blackwell Publishers.

Mackie, J. (1974). *The cement of the universe*. Oxford: Clarendon Press.

Mackie, J. (1980). The transitivity of counterfactuals and causation. *Analysis Oxford*, *40*(1), 53–54.

McDermott, M. (1995). Redundant causation. *The British Journal for the Philosophy of Science*, *46*(4), 523–544.

Menzies, P. (2009). Counterfactual theories of causation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2009 ed.).

Merricks, T. (2001). *Objects and persons*. Oxford: Clarendon Press.

Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.

Sartorio, C. (2006). Failures to act and failures of additivity. *Philosophical Perspectives*, *20*, 373–385.

Schaffer, J. (2000). Trumping preemption. *The Journal of Philosophy*, *97*(4), 165–181.

Schaffer, J. (2003a). Is there a fundamental level? *Noûs*, *37*(3), 498–517.

Schaffer, J. (2003b). Overdetermining causes. *Philosophical Studies*, *114*(1–2), 23–45.

Woodward, J. (2016). The problem of variable choice. *Synthese*, *193*(4), 1047–1072.