REGULARITY THEORIES REASSESSED

Michael Baumgartner
University of Bern

Abstract

Until very recently, regularity accounts of causation have virtually vanished from the scene. Problems encountered within other theoretical frameworks have lately induced philosophers working on causation – as e.g. Hall (2004) or Beebee (2006) – to direct their attention back to regularity theoretic analyses. In light of the most recent proposals of regularity theories, the essay at hand will therefore reassess the criticism brought forward against regularity accounts since Mackie's famous, yet failed, (1974) attempts at analyzing causation with recourse to regularities among types of events. It will be shown that most of these objections target strikingly over-simplified regularity theoretic sketches, which no present-day regularity theorists would seriously consider worth a second thought. By outlining ways to refute these objections it will be argued that the prevalent conviction as to the overall failure of regularity theories has been hasty – to say the least.

1 Introduction

A mere glance at the abundance of controversial literature on causation, published during the past 30 years, reveals that regularity accounts of causation – until very recently – virtually vanished from the scene. For lack of space and interest, studies not primarily concerned with causation every now and then roughly explicated our causal intuitions in terms of regularities, but hardly anybody seriously wanting to analyze causation resorted to regularity accounts any more. Problems encountered within other theoretical frameworks have lately induced philosophers working on causation – as e.g. Hall (2004) or Beebee (2006) – to direct their attention back to regularity theoretic analyses. In light of the most recent proposals of regularity theories as can be found in Graßhoff and May (2001) and May (1999), the essay at hand will therefore reassess the criticism brought forward against regularity accounts since Mackie's famous, yet failed, (1974) attempts at analyzing causation with recourse to regularities among types of events.

Notwithstanding the scepticism encountered by Hume's regularity theoretic successors, there are several commonly acknowledged advantages of an analysis of causation in terms of regularities. A regularity theoretic notion of causation directly mirrors central pre-theoretic intuitions with respect to the cause-effect relation expressible in well-known principles as "The same cause is always accompanied by the same effect" or "If no cause is present, no effect occurs". Moreover, the conceptual apparatus resorted to by a regularity theoretic analysis is fully embedded within the uncontroversial and well mastered area of extensional standard logic. Furthermore, unlike e.g. counterfactual accounts regularity theories straightforwardly handle cases of overdetermination. As against interventionist or manipulatory accounts, analyses of causation in terms of regularities do not run the risk of being anthropocentric. Contrary to probabilistic accounts, regularity theories are not compromised by paradoxical data as, for instance, generated in cases of Simpson's Paradox. Finally, while transference theories treat a fundamental type of causal process, transference processes, as conceptually primitive and thus do not attempt to provide a reductive analysis of causation, regularity accounts, properly conceived, offer the promising prospect of explicating the cause-effect relation – or at least one among possibly several types of cause-effect relations[1] – in entirely non-causal terms.

Due to limited space, a regularity theoretic analysis of causation cannot be fully developed here. I have done this elsewhere.[2] The pages to come will thence simply review the objections that traditionally have been raised against regularity accounts. It will be shown that most of these objections target strikingly over-simplified regularity theoretic sketches, which no present-day regularity theorists would seriously consider worth a second thought. By outlining ways to refute these objections it will be argued that the prevalent conviction as to the overall failure of regularity theories has been hasty – to say the least.

2 Hume's Legacy

The philosophical core of regularity theories of causation, concisely put, consists of three main tenets: (i) anti-realism with respect to the ontological status of the causal relation, (ii) general causation[3] – causation on type-level – as primary analysandum, and (iii) universal regularities among event types as primary analysans.

According to Hume, the godfather of regularity theories, single event sequences are not identifiable as being of causal nature by some inherent physical feature or property. A causal interpretation of an event sequence is warranted only if the corresponding events, understood as spatiotemporally located tokens or particulars,[4] instantiate factors or event types[5] which satisfy a material conditional as "Whenever A is instantiated, B is instantiated" such that the instances of A and B differ and are spatiotemporally proximate.[6] Events do not cause themselves – no self-causation – and effects occur nearby their causes – no action at a distance. For an event a to be identified as a cause of another event b, it is, according to this conception, required that a instantiates a factor A whose instances are always followed by events of type B, which is instantiated by b. Causes are thus analyzed to be *sufficient conditions* of their effects. This yields a first Hume-inspired proposal for a regularity theoretic account of causation:

(I) A is a cause of B iff A is sufficient for B and the instances of A and B differ and are spatiotemporally proximate.

(I), as well as the other regularity theoretic sketches discussed in this paper, heavily relies on the notions of differing instances of factors and of spatiotemporal proximity. Both of these notions call for clarifications, which for lack of space, however, cannot be provided here. A mere pre-theoretic and intuitive understanding of when causal relata are identical as opposed to different and proximate as opposed to distant satisfactorily meets the requirements of the upcoming considerations. Moreover, note that the relational constraints (I) imposes on the instances of causes and effects forestall an explication of the notion of a sufficient condition in terms of a simple propositional material conditional. The same holds for necessary conditions, which will become of importance below. Sidestepping all technical details, the notions of sufficient and necessary conditions thus have to be understood with recourse to the first-order formalism. Broad (1930) proposes the following definitions, which for reasons of simplicity will not be explicitly (formally) spelled out here:

> "C is a *sufficient condition* (…) of E" means "Everything that has C has E".
> "C is a *necessary condition* (…) of E" means "Everything that has E has C".[7]

The simple identification of causes and sufficient conditions that are instantiated nearby their conditioned factors as exemplified in (I) is the target of many well known criticisms of regularity accounts.

3 Imperfect Regularities

There is an obvious first objection to this overly simple variety of a regularity analysis: Most causes plainly are not sufficient for their effects in the sense of (I) or, as Hitchcock (2002) puts it, causal regularities commonly are *imperfect regularities*. It is even highly dubious whether there in fact are any universal regularities as required by (I) in nature at all. A factor A being a sufficient condition of a factor B such that the (differing) instances of A and B are spatiotemporally proximate, thus, clearly is not a necessary condition of A causing B.

Hume did not ignore the fact that factors whose instances are not universally correlated may nonetheless be causally dependent. In this respect, (I) does not fully reproduce Hume's analysis. In order for a factor A to be identifiable as a cause of a factor B, Hume did not require A to be sufficient for B simpliciter, i.e. sufficient relative to any causal background. Rather, he devised a cause to be sufficient for its effect only when "plac'd in like circumstances".[8] Hence, common analyses of causation in terms of sufficient conditions are supplemented by a *ceteris paribus clause* such that causes are merely required to be ceteris paribus sufficient for their effects. These considerations induce a modification of (I) to:

(II) A is a cause of B iff A is ceteris paribus sufficient for B and the instances of A and B differ and are spatiotemporally proximate.

However, the notion of a ceteris paribus clause is notoriously vague. (II) is only fruitfully applicable given a proper explication of the ceteris paribus proviso. For our present purposes, a rough idea of what traditionally is meant by a ceteris paribus clause will suffice. In order to provide such a rough idea, consider a match that is struck against a matchbox and that, as a consequence thereof, catches fire. Refer to this scenario as "S_1". What requirements does a second scenario S_2 have to satisfy such that Hume would classify S_1 and S_2 as "like circumstances"? Obviously, no two scenarios coincide with respect to all their properties or characteristics. At least in its spatiotemporal properties any scenario (or circumstance) S_2 diverges from S_1. For Hume to speak of "like circumstances" S_1 and S_2 only have to share certain significant properties. The match in S_1 is struck with a certain speed and thrust, it is exposed to a certain amount of friction, and its flammable head is dry. Moreover, S_1 features the presence of enough oxygen. If S_2 coincides with S_1 relative to these kinds of properties, only a proper subset of which have been explicitly included in the above list, S_2 can be said to satisfy the ceteris paribus clause with respect to S_1, i.e. S_1 and S_2 can be referred to as "like circumstances". The properties that have to coincide in like circumstances share a common feature: They are all causally relevant to the effect under consideration. This induces a modification of (I) and (II) to:

(III) A is a cause of B iff there is a scenario S_1 such that A is sufficient for B in S_1 as well as in all scenarios that agree with S_1 as regards all causally relevant features and the instances of A and B differ and are spatiotemporally proximate.

Explicating the ceteris paribus proviso in this vein, of course, gives immediate rise to circularity objections.[9] (III) cannot be considered an analysis of the basal causal notion any longer, for the definiens itself presupposes the notion of causal relevance. In order to determine relative to which features of background conditions a factor A is sufficient for a factor B, (III) calls for clarity on the causes of B, which is just what (III) pretends to provide at the same time. Hence, integrating an explication of the ceteris paribus proviso along the lines of (III) into an analysis of causation is not feasible. Nonetheless, the above considerations reveal an important feature of causal dependencies: They are not one-to-one, but many-to-one dependencies. Or put differently, while effects correspond to single factors, causes are complexes of *jointly* instantiated factors. Consequently, striking a match is not autonomously sufficient for the match to catch fire. Rather, factors as striking a match with a certain speed and thrust, dryness of its flammable head, presence of enough oxygen etc. are *jointly* sufficient for the match to light. Among the instances of these factors very specific spatiotemporal relations must subsist in order for their combination to actually become causally sufficient. Bypassing the problem of clarifying these spatiotemporal constraints for now, the ceteris paribus proviso can now be suitably accounted for without explicitly having to integrate it into an analysis of causation. For if a cause is no longer held to be autonomously sufficient for its effect, but is taken to be *a mere part* of a sufficient condition, the ceteris paribus clause can be dropped from (II) and (III) respectively.

(IV) A is a cause of B iff A is a part of a sufficient condition of B and the instances of A and B differ and are spatiotemporally proximate.

As long as sufficient conditions are simply understood to be antecedents of (universally quantified) conditionals as indicated in section 2, the notion of *a part* of a sufficient condition is straightforwardly explicable in terms of *conjuncts* of such antecedents. (IV) is not refuted by a struck match that does not catch fire. Whenever a match is struck, but fails to light, it may now be argued that – notwithstanding the striking – not all factors of the corresponding complex sufficient condition for lighting matches have been instantiated on the respective occasion.

4 Monotony

Implementing regularities along the lines of (IV) to identify causal dependencies still does not amount to a feasible analysis of causation, because there are regularities of the required type that are not amenable to a causal interpretation. One such type of regularities is due to the law of monotony: Antecedents of conditionals can *salva veritate* be conjunctively supplemented by further factors. Monotony allows for arbitrarily constructing complex sufficient conditions that are by no means causally interpretable. Consider again the match example discussed above. Striking a match with a certain speed and thrust, factor A, dryness of its flammable head (B), and presence of enough oxygen (C), shall be assumed to be jointly sufficient for the corresponding match to catch fire (D), i.e. $A \wedge B \wedge C \rightarrow D$. Yet, if A, B, and C are jointly sufficient for the match to light, the combination of $A \wedge B \wedge C$ and singing a song is thus sufficient, too. Moreover, $A \wedge B \wedge C$ combined with singing a song and blinking an eye and wiggling the left pinky toe are also going to be jointly sufficient for the match to catch fire. Or formally:

$$A \wedge B \wedge C \rightarrow D \vdash A \wedge B \wedge C \wedge X \rightarrow D, \tag{1}$$

where X stands for an arbitrary factor or conjunction of factors. This demonstrates that being a part of a sufficient condition, i.e. being a conjunct within a sufficient conjunction of factors, is by no means sufficient for being a cause of the respective conditioned factor.

Broad (1930) has been the first to propose a solution to this problem. He does not analyze causes to be mere parts of sufficient conditions, but rather to be *non-redundant* parts of such conditions. A non-redundant part of a sufficient condition can be spelled out – in purely logical terms[10] – as being a conjunct of a sufficient condition such that if it is eliminated from that condition the latter loses its sufficiency for a corresponding effect. Complex causes then are no longer understood as conjunctions of factors which are jointly merely sufficient for their effect, but are newly taken to be *minimally sufficient* conjunctions of factors – a minimally sufficient conjunction being a conjunction that does not have sufficient proper parts.

(V) A is a cause of B iff A is a part of a minimally sufficient condition of B and the instances of A and B differ and are spatiotemporally proximate.

Applying (V) to the match example prohibits a causal interpretation of, say, the combination of striking a match, presence of enough oxygen, dryness of the match, and singing a song. The conjunction of these factors is merely sufficient, but not minimally sufficient for the match to catch fire. One of its conjuncts, the singing, can be eliminated without loss of sufficiency. Requiring a minimalization of sufficient conditions in Broad's sense precludes a causal interpretation of arbitrary extensions of sufficient conditions based on the law of monotony.

5 Empty Regularities

Minimalizing sufficient conditions does not solve all the problems that can be induced by monotony. Consider again the match example. As we have seen, the presence of oxygen, factor C, is not itself sufficient for a match to catch fire (D). Other factors – A and B – have to be instantiated as well in order for D to occur. Yet, $A \wedge B \wedge C$ is not the only minimally sufficient condition containing C. Another such condition is constituted by the presence of oxygen and the absence of oxygen: $C \wedge \neg C$. A contradiction is sufficient for any factor, not only for matches catching fire, but also for rain to fall and elephants to be born. That $C \wedge \neg C$ is moreover minimally sufficient for a match to light can easily be verified by either removing C or $\neg C$, both of which is accompanied by a loss of sufficiency for D. More generally put: Material conditionals are true if their antecedents are false or non-instantiated, or *empty* for short. Any regularity statement along the lines of (V) whose antecedent is empty is, accordingly, termed an *empty regularity*. Empty regularities do not only result from contradictory antecedents and, thus, from logically non-instantiatable antecedents, but also from physically non-instantiatable antecedents as, for instance, "Whenever Pegasus goes skiing, Lake Thun is made of gold".

The truth of empty regularity statements raises another often cited problem for regularity accounts: Empty regularities are, notwithstanding their truth, not amenable to a causal interpretation.[11] The combination of absence and presence of oxygen – $C \wedge \neg C$ – does not cause the sinking of Mississippi steamers, even though $C \wedge \neg C$ in fact is minimally sufficient for these sinkings. Neither can Pegasus' ski tour be seen as a cause of the golden content of a lake.

Solutions to this problem are easily thought of. It is not the case that only a certain proper subset of all empty regularities consists of regularities that are not causally interpretable, rather, *no* empty regularities are thus interpretable. Causal dependencies subsist among entities that exist in nature. Inexistent things may not be causally related.[12] Therefore, empty regularities can straightforwardly be excluded from causal interpretability by adding a further constraint to (V) that requires the antecedent of causally interpretable regularities to be non-empty.

(VI) A is a cause of B iff the following conditions hold:

- (i) A is a part of a minimally sufficient condition X_1 of B,
- (ii) the instances of A and B differ and are spatiotemporally proximate, and
- (iii) there is an instance of X_1.

Alternatively it might be argued that event types or conjunctions of event types without instances must not even be taken into consideration in the first place when it comes to causal analyses. Thus, existence requirements with respect to causally analyzed factors might be imposed as a kind of criterion of well-formedness for causal factors. Upon opting for this solution to the empty regularities problem, which essentially amounts to the same as a solution in the vein of (VI), an explicit modification of (V) can even be dispensed with by simply relativizing regularities of type (V) to well-formed causal factors.

6 Asymmetry

Another objection often raised against regularity accounts concerns the asymmetry of causation. The cause-effect relation is asymmetric. A being a cause of B neither implies B to be a cause of A nor $\neg B$ to be a cause of $\neg A$. However, by contraposition $A \rightarrow B$ is a true material conditional iff $\neg B \rightarrow \neg A$ is so too. Which of these true conditionals is to be causally interpreted? It is certainly not the case that a factor is causally relevant to another factor iff the negation of the latter is causally relevant to the negation of the former. Smoking is causally relevant to lung cancer, without the absence of lung cancer having causal impact on abstinence from smoking. Accordingly, many critics of regularity theories have claimed that regularity accounts cannot adequately distinguish between causes and effects.[13] (VI) identifies A to be cause of B iff it identifies $\neg B$ to be cause of $\neg A$, which indicates that we have not come up with an adequate regularity theoretic analysis of causation yet.

Satisfactorily mirroring the asymmetry of causation is an intricate problem not only faced by regularity accounts, but by virtually all presently known theories of causation. Normally the direction of causation is accounted for with recourse to some asymmetry external to the conceptual framework used in the analysans of causation as – most prominently – the direction of time or human manipulation and intervention.[14] Applied to the regularity theory considered here, this could possibly yield that material conditionals in the sense of (VI) are causally interpretable only if the instances of the antecedent *precede* the instances of the consequent. Along these lines, one of $A \rightarrow B$ and $\neg B \rightarrow \neg A$ could be excluded from causal interpretability. However, accounting for the asymmetry of causation by means of an asymmetry that is *external* to the conceptual framework of a respective analysans of the causal relation is a high theoretical price to pay. One would have to account for the direction of time independently of the direction of causation and thus deviate from an often adopted programme in the philosophy of time that takes the direction of causation to be primary.[15] Resorting to manipulation, on the other hand, relativizes the asymmetry of causation to human intervention, where intuitively this asymmetry seems to be perfectly independent of human existence. Causal processes – as planetary movements or volcanic eruptions on Saturn – that are not manipulable by humans are asymmetric and orientable just as everyday earthly processes as the breaking of a window or the starting of a car engine which are open to human intervention. Moreover, it is unclear how the notions of agency, intervention, and manipulation could be clarified without recourse to the causal relation. In fact, these notions seem to straight-out presuppose clarity on causation.

This is a generalizable consequence of implementing any external asymmetry for a theoretical account of the asymmetry of causation: The external asymmetry becomes more basic than the cause-effect relation. Thereby a straightforward causal analysis of these external asymmetries is blocked. However, intervention, for instance, can hardly be more transparently analyzed than in terms of causal processes where human action is involved as a cause. As long as we are not inevitably constrained to an analysis of the direction of causation by means of an external asymmetry, theoretical foresight calls for abstinence from recourse to such asymmetries. Indeed, regularity theories are capable of capturing the direction of causation without recourse to asymmetries that are external to the conceptual framework of a regularity theoretic analysans of causation.

The cause-effect relation can be oriented on mere logical grounds. Roughly, while conditional dependencies among single factors cannot be attributed a direction without resorting to external asymmetries, complex *nets* of such dependencies are orientable based on existing regularities only. There are several alternative causes for each effect.

A match can be lit by either striking it against a match box, by exposing it to fire or to a flammable chemical etc. Accordingly, causally interpretable regularities are far more complex than expressed by (VI). Rather than merely one minimally sufficient condition $A \wedge C \wedge D$, a whole number of alternative minimally sufficient conditions – $A \wedge C \wedge D$, $E \wedge F \wedge G$, $H \wedge I \wedge J$,... – must be invoked for each effect. On the other hand, an effect does not occur without the presence of at least one of its alternative causes. Thus, whenever the effect is given, at least one of its alternative minimally sufficient conditions is given as well. These mutual dependencies among causes and effects are tentatively[16] expressible by means of a biconditional as in (2).

$$(A \wedge C \wedge D) \vee (E \wedge F \wedge G) \vee (H \wedge I \wedge J) \leftrightarrow B \tag{2}$$

Each complex cause of B is minimally sufficient for B, while the disjunction of all alternative causes is *necessary* for B.[17] (2) is not symmetrical with respect to the factors to the left and the right of "$\leftrightarrow$". The instantiation of a particular disjunct is minimally sufficient for B, but not vice versa. B does not determine a particular disjunct to be instantiated.[18] B only determines the whole disjunction of minimally sufficient conditions. Hence, given that an instantiation of $A \wedge C \wedge D$ is observed, it can be inferred that there is an instance B somewhere in the corresponding spatiotemporal neighborhood. On the other hand, if an instance of B is observed, no such inference to a proximate instantiation of $A \wedge C \wedge D$ is possible. The observed instance of B might well have been caused by $E \wedge F \wedge G$. This asymmetry corresponds to the asymmetry of determination. It induces a specification of (VI) along the following lines:

(VII) A is a cause of B iff the following conditions hold:

- (i) A is a part of a minimally sufficient condition X_1 of B,
- (ii) X_1 is a disjunct contained in a disjunction $X_1 \vee X_2 \vee \ldots \vee X_n$, $n \geq 2$, of other minimally sufficient conditions of B, such that $X_1 \vee X_2 \vee \ldots \vee X_n$ is necessary for B,
- (iii) the instances of A and B differ and are spatiotemporally proximate, and
- (iv) there is an instance of X_1, X_2, ..., and of X_n.

Clearly though, by contraposition (2) is equivalent to

$$\neg B \leftrightarrow \neg(A \wedge C \wedge D) \wedge \neg(E \wedge F \wedge G) \wedge \neg(H \wedge I \wedge J) \tag{3}$$

However, in view of the fact that effects have several alternative causes, (VII) restricts the causal interpretability of complex regularity statements to one specific syntactical form. Within a set of logically equivalent regularity statements, only expressions with a syntax that exhibits alternative minimally sufficient conditions as *disjuncts* of a necessary condition are causally interpretable. Applied to (2) and (3), this syntactical constraint prohibits a causal interpretation of (3) for it does not render an underlying causal structure transparent in the sense just delineated.

(3) is moreover equivalent to a biconditional that results from (3) by factoring out and bringing the righthand side back into disjunctive normal form:

$$(\neg A \wedge \neg E \wedge \neg H) \vee (\neg A \wedge \neg E \wedge \neg I) \vee \ldots \vee (\neg D \wedge \neg G \wedge \neg J) \leftrightarrow \neg B \tag{4}$$

In contrast to (3), (4) is unproblematically causally interpretable. While (2) identifies three minimally sufficient conditions as complex causes of B, (4) establishes the causally interpretable minimally sufficient conditions of $\neg B$. Each of those conditions amounts to a conjunction consisting of the negation of exactly one conjunct of each disjunct of (2). Furthermore, (4) does not reverse the direction of the causal dependencies expressed in (2). Both identify B and $\neg B$ respectively as effects and the other factors as causes. Thus, there is one regularity statement complying to the syntactical constraints imposed by (VII) for a positive effect and one for the latter's negative complement. Both of these regularities exhibit the same asymmetry. Accordingly, neither of them poses a problem for (VII).

Accounting for the asymmetry of causation in this vein has an important implication as regards the minimal complexity of causal structures. A factor or conjunction of factors X_1, that is both minimally sufficient and necessary for another factor or conjunction of factors X_2, cannot be identified as cause of X_2, for X_2 would be minimally sufficient and necessary for X_1 as well. All empirical evidence such a dependency structure would generate are perfectly correlated instantiations of X_1 and X_2 – both would either be co-instantiated or absent. Such empirical data is not causally interpretable. In order to distinguish causes from effects and to orient the cause-effect relation, at least *two* alternative causes are needed for each effect.

Plainly, this is merely a rough sketch of how the direction of causation can be accounted for on regularity theoretic grounds. More would have to be said on these matters. For now, however, it must suffice to note that, contrary to the widespread opinion in the literature, regularity theories not only seem capable of adequately capturing the asymmetry of the cause-effect relation, but moreover offer the prospect of successfully doing so without resorting to asymmetries external to the conceptual framework implemented in their analysans of causation. Against this background, such external asymmetries as the direction of time or of human intervention remain amenable to a straightforward analysis in terms of the asymmetry of causation.

7 *Spurious Regularities*

One of the most widespread criticisms against regularity theories stems from so-called *spurious regularities*.[19] Consider two parallel effects A and B of a common cause C – a structure commonly labelled an *epiphenomenon* – and assume for simplicity's sake that C in fact is minimally sufficient for A and B. Such as to do justice to the complexity of causal structures let us suppose there exists one minimally sufficient alternative cause for A and B each – D for A and E for B. All in all, the causal structure under consideration thus is assumed to be of a form as depicted in figure 1.[20] In this constellation, A in combination with the absence of D, i.e. $A \wedge \neg D$, is minimally sufficient for B without $A \wedge \neg D$ being a complex cause of B. Whenever $A \wedge \neg D$ occurs, C is present as well, for no effect occurs without any of its causes. Hence, if D is absent, C must be present to account for A. Furthermore, since C is taken to be sufficient for B, it follows that $A \wedge \neg D$ is thus sufficient as well. Of course, $A \wedge \neg D$ is moreover part of a necessary condition of B:

$$(A \wedge \neg D) \vee C \vee E \leftrightarrow B \tag{5}$$

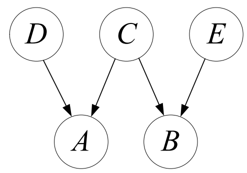

Fig. 1: An epiphenomenon that gives rise to spurious regularities.

According to (VII), (5) is a regularity statement that is causally interpretable. This clearly is an unacceptable consequence, for, as mentioned above, relative to the construction of the structure in figure 1, $A \wedge \neg D$ does not cause B.

Structures as the one under consideration are ubiquitous in nature. The most famous concrete example of this type, undoubtedly, is the so-called *Manchester-Factory-Hooters* example based on which Mackie (1974) ultimately abandoned the attempt to provide a genuine regularity theoretic analysis of causation.[21] Examples of this type unmistakably demonstrate that necessary conditions, just as sufficient conditions, may contain redundant factors. A being necessary for B implies that $A \vee C$ is necessary for B. Or formally:

$$B \rightarrow A \vdash B \rightarrow A \vee C. \tag{6}$$

Any true conditional stays true if any (true or false) disjunct is added to its consequent. Analogous to the case of sufficient conditions, the extendability of necessary conditions by arbitrary disjuncts forecloses a causal interpretability of necessary conditions. A causal interpretation of necessary conditions is only warranted if the conditions exclusively contain factors that are essential to the bringing about of the purported effect. Arbitrary factors as C in (6) or conditions as $A \wedge \neg D$ in (5) must, even if they are minimally sufficient, not be incorporated in causally interpretable necessary conditions.

Graßhoff and May (2001) and May (1999) have proposed an analogous solution to this problem as in case of the difficulties induced by monotony. They call for a minimalization of necessary conditions. The basic idea behind the minimalization of necessary conditions coincides with the criterion guiding the minimalization of sufficient conditions: A necessary condition is *minimally* necessary iff it does not contain a necessary proper part. Minimalizing necessary conditions based on this notion of a minimally necessary condition in fact eliminates all and just the spurious minimally sufficient conditions as $A \wedge \neg D$ from complex regularity statements as (5). Properly substantiating this point, however, requires a certain technical apparatus that cannot be introduced here.[22] For now, it must suffice to illustrate the minimalization of (5). Consider again the structure depicted in figure 1: Whenever B is given, either C or E is instantiated. Thus, $C \vee E$ is necessary for B. The antecedent of (5) has no other necessary proper part. $(A \wedge \neg D) \vee C$ is not necessary, for there are instances of B without $A \wedge \neg D$ and C being instantiated – say, when $A \wedge D$ is given along with $\neg C$ and E. Neither is $(A \wedge \neg D) \vee E$ necessary for B: There are instances of B without instances of $A \wedge \neg D$ and E occurring – for example, when $A \wedge D$ is given in combination with $\neg E$ and C. Thus, there in fact exists a suitable refinement of (VII) that allows for an appropriate handling of spurious correlations on regularity theoretic grounds.

(VIII) A is a cause of B iff the following conditions hold:

(i) A is a part of a minimally sufficient condition X_1 of B,
(ii) X_1 is a disjunct contained in a disjunction $X_1 \vee X_2 \vee \ldots \vee X_n$, $n \geq 2$, of other minimally sufficient conditions of B, such that $X_1 \vee X_2 \vee \ldots \vee X_n$ is *minimally* necessary for B,
(iii) the instances of A and B differ and are spatiotemporally proximate.

As May (1999) has shown, introducing a minimality constraint on necessary conditions has a very important existential implication that allows for dropping condition (iv) of (VII). $C \vee E$ being minimally necessary for B implies there being an instance of B without a corresponding instance of C – refer to this scenario by S_1 – and an instance of B without a corresponding instance of E – scenario S_2. That is, none of the two disjuncts is itself necessary for B. Nonetheless the disjunction as a whole is thus necessary. Therefore, both in S_1 and in S_2 there must be an instance of $C \vee E$. Since in S_1, by assumption, there is no event of type C, there must be an event of type E – and vice versa for S_2. This finding is generalizable: Every disjunct of a minimally necessary condition is instantiated at least once along with the corresponding effect, when *all* the other disjuncts are absent.[23]

8 Single-Case Regularities

A problem that is closely related to the problem of empty regularities has e.g. been raised by Armstrong (1983). A conditional turns out true if both its antecedent and consequent are true. Thus, if antecedent and consequent of a conditional each report the occurrence of a singular event that actually has occurred, the conditional as a whole is true. Therefore, Armstrong argues, a regularity as required by a regularity theory subsists among any two occurring events, irrespective of whether they are causally related or not. No doubt, a conditional as "Whenever Nero sets fire on Rome, the Titanic sinks" is true and no doubt, we are not prepared to hold Nero causally responsible for the sinking of the Titanic. Hence, Armstrong's argument continues, not only empty, but also these so-called *single-case* regularities pose a serious problem for a regularity theoretic analysis of causation.

At first, it must be pointed out that the plain truth of a conditional as "Whenever Nero sets fire on Rome, the Titanic sinks" does not suffice to identify Nero's setting fire on Rome as a cause of the sinking of the Titanic according to any of the regularity theoretic accounts (I) to (VIII) considered thus far. For these accounts not only require causes and effects to satisfy a material conditional, but moreover to be proximately instantiated. Even though the notion of spatiotemporal proximity has not been properly explicated here, relative to any pre-theoretic understanding of that notion, it seems plain that Nero's setting fire on Rome and the sinking of the Titanic cannot be seen as proximate events. However, this shortcoming of Armstrong's argument is easily remedied. Assume that Harold Bride, the junior wireless operator on the Titanic, for the first (and only) time in his life lit a Havana cigar moments before the ship hit the iceberg. The conditional "Whenever Harold Bride lights a Havana, the Titanic sinks" is true and, moreover, the events mentioned in its antecedent and consequent are spatiotemporally proximate. Of course, Bride's lighting of a Havana is not only sufficient, but moreover minimally sufficient for the Titanic to sink. The antecedent of the above mentioned conditional does not comprise a sufficient proper part. Furthermore, Bride's lighting of a Havana is not the only minimally sufficient condition of the sinking of the Titanic. The latter's real cause constituted by the collision with the iceberg amounts to another such condition. Hence, there is a necessary condition of the sinking of the Titanic that contains Bride's lighting of a Havana as a minimally sufficient disjunct. This suffices to refine Armstrong's argument such that it does justice to the complexity of causal structures as required by (VII): Any two actually occurring events that are spatiotemporally proximate satisfy a regularity as required by (VII), yet by no means all thus related events are causally related as well. Consequently, (VII) does not amount to a sufficient condition for causal relatedness.

That (VII) is unsuited as analysans of causation has already been demonstrated by the problem of spurious regularities. In order for Armstrong to succeed in establishing that single-case regularities prove the fundamental defectiveness of regularity accounts, his argument must be tailored to be directed against (VIII). It thus must be shown that Bride's lighting of a Havana is not only contained in a necessary condition of the sinking of the Titanic, but is moreover a non-redundant part of a causally interpretable minimally necessary condition thereof.

Before this further refinement will be attempted a possible objection against Armstrong's argument has to be considered. Antecedent and consequent of "Whenever Harold Bride lights a Havana cigar, the Titanic sinks" involve proper names or, if formal explications by means of definite descriptions are preferred, predicates that apply to single events only – more generally: *local predicates*, i.e. predicates that involve spacetime coordinates or singular terms. The admissibility of local predicates in law-like contexts, as is well known, is commonly denied in the literature. Causal dependencies do not exclusively subsist in local domains as the one constituted by the Titanic's maiden voyage. Harold Bride's lighting of a Havana cigar is a cause of the sinking of the Titanic if and only if lighting Havana cigars generally cause ocean liners to sink. In view of this universality of causal dependencies and, moreover, in light of the basic intuition behind regularity accounts according to which only repeated instantiations of factors allow for causal diagnoses, it is plain that proper names and local predicates must be excluded from well-formed factor definitions. Armstrong's argument might thus be rejected for its involvement of factors that are defined by means of local predicates and that, accordingly, are not well-defined causal factors. Yet, as Armstrong points out, especially in macroscopic contexts as the one under consideration every predicate involving local constraints may well be replaceable by a co-extensional non-local predicate.[24] For instance, "…is Harold Bride's lighting of a Havana cigar" could be replaced by a conjunction of arbitrary non-local properties that, taken together, happen to apply to exactly one event, namely Harold Bride's lighting of a Havana cigar. Or instead of by use of a proper name, Harold Bride might be referred to by specifying his genome, while the Titanic is individuatable by means of its molecular structure. So let us grant that "Whenever Harold Bride lights a Havana cigar, the Titanic sinks" represents the exact same single-case regularity as "Whenever a person with genome s lights a Havana cigar, an ocean liner with molecular structure t sinks", or formally $S \rightarrow T$, which constitutes a single-case regularity involving non-local predicates and thus well-defined causal factors only.

The Titanic's collision with the iceberg, of course, is expressible by means of non-local predicates as well. Symbolizing this (non-locally defined) collision by C we get the following true biconditional that is causally interpretable according to (VII):

$$S \vee C \leftrightarrow T. \tag{7}$$

If $S \vee C$ is not merely necessary, but moreover *minimally* necessary for T, (7) not only refutes (VII), but also (VIII). $S \vee C$, however, does not amount to a minimally necessary condition of T. There is only one single instance of each factor involved in (7). Whenever T occurs, both S and C are present nearby. Thus, the antecedent of (7) can be further minimalized:

$$S \leftrightarrow T \tag{8}$$
$$C \leftrightarrow T \tag{9}$$

Neither (8) nor (9), however, are causally interpretable, for these expressions merely report a perfect correlation of T and S and C respectively. Any of these factors is given if and only if the other two factors are given as well. Such perfect correlations, as we have seen above, are not causally interpretable, for none of the involved factors is identifiable as cause and effect respectively. Since (VIII) requires causally interpretable regularities to specify minimally necessary conditions of a certain minimal complexity, neither (8) nor (9) is amenable to a causal interpretation according to (VIII). That, however, does not mean that (VIII) does not allow for identifying the collision with the iceberg as a cause of the sinking of the Titanic. A different and more coarse-grained typing of the occurrences involved in the sinking of the Titanic will yield far more instances for each causal factor, which, in turn will suspend biconditional dependencies as in (8) or (9).

All in all, a regularity theoretic analysis of causation imposes certain minimal complexity constraints on causal structures. Causal structures are not one-to-one dependencies among single factors. Every effect has several alternative complex causes. In order to establish such dependencies, more than one single instance of causes and effects is required. The regularity theoretic notion of causation expressed in (VIII) mirrors these minimal complexity requirements and, accordingly, is not affected by the problem of single-case regularities.

9 Singular Causation

The problem posed by single-case regularities demonstrates that a regularity theory cannot be successful if its analysans is taken to be singular causation, i.e. causation among token events. In order to account for the complexity of a causal structure several instances of that structure are needed. Nonetheless, some regularity theories, as e.g. developed in Mackie (1965), primarily analyze singular causation. A lot of the critisicm raised against regularity theories over the past three decades targets this kind of singularist account. Some exemplary objections include the following: Among others, Collins, Hall, and Paul (2004) claim, singularist regularity theories cannot adequately handle cases of preemption. Kim (1973) criticizes Mackie (1965) for not being able to adequately assign token events as causes to concrete singular effects – a problem that has come to be known as the *pairing problem*. Finally, many authors – good examples of which can also be found in Collins, Hall, and Paul (2004) – have raised doubts about whether regularity theories can satisfactorily handle cases of causation among absences and omissions.

First of all, none of these objections exclusively aims at regularity theories. Counterfactual causation – as is well known – faces fundamental difficulties when it comes to cases of preemption.[25] The pairing problem, in turn, affects any account of causation that imposes some (deliberately) vague spatiotemporal proximity constraints on the causal relata. And any analysis of causation that takes the causal relata to be events has a hard time accounting for causation among absences and omissions. Second, since my main concern here is the canonical criticism exclusively brought forward against regularity theories of the Humean type, this singularist branch of criticism cannot not be given adequate consideration in the present context.

Nonetheless, however, any philosophical account of causation that wants to be taken seriously has to say something about singular causation. Thus, a regularity theory cannot contend itself with analyzing general causation. Any theory that is primarily interested in general causation has to provide some account of how causal dependencies among tokens are derivable from causal dependencies among types. In light of (VIII), the following transition from general to singular causation is easily thought of:

Singular Causation: Event a is a cause of event b iff the following conditions hold:

(i) a instantiates a factor A and b instantiates a factor B such that A and B are causally related in terms of (VIII)
(ii) every other factor X, that is part of that minimally sufficient condition of B which A is part of, is instantiated coincidently with a.

To what extent this analysis of singular causation successfully deals with cases of preemption, the pairing problem, and with causation among absences and omissions has to left open here. What matters for now is that there are no principled obstacles to implementing an analysis of general causation along the lines of (VIII) when it comes to accounting for singular causation.

10 Indeterminism

The discussion of the thus far considered arguments against regularity accounts revealed that common objections either attack overly simplistic or singularist regularity theoretic analyses. None of these objections affects a regularity account that both adequately represents the whole complexity of causal structures and gives preference to type-level dependencies. Still, there is one conventional argument often put forward against regularity theories that neither targets an oversimplified nor a singularist account. With the advent of quantum mechanics, it appears that some processes as e.g. radioactive decay both run irreducibly indeterministically and are to be qualified as being of causal nature. Accordingly, the second half of the 20$^{\text{th}}$ century has seen the rise of probabilistic analyses of causation as for instance proposed by Reichenbach (1956) or Suppes (1970), who took these indeterministic processes as sufficient evidence to the effect that building a theory of causation on the principle of determinism, as regularity accounts generally do, is fundamentally mistaken.

While the previously considered objections do not question the principle of determinism and thereby the conceptual fundament of regularity accounts, that is just what this argument from indeterminism does. Criticizing regularity theories in this vein presupposes that there are causal processes that run irreducibly indeterministically. While the existence of irreducibly indeterministic processes is hardly challengeable according to standard interpretations of quantum mechanics[26], there are many open questions – as for instance raised by phenomena of the EPR type[27] – with respect to the causal interpretability of these processes. The causal interpretability of indeterministic processes is even more questionable as not even modern probabilistic theories of causation can successfully account for these processes. The latter violate central assumptions of probabilistic analyses of the cause-effect relation, as the Reichenbachian common cause principle[28] or the causal Markov assumption.[29] In consequence thereof, probabilistic accounts of causation are forced to limit their scope to so-called pseudo-indeterministic processes, i.e. processes whose indeterminacy is merely due to incomplete knowledge of or control over the involved factors.[30]

The question as to whether the irreducibly probabilistic progression of quantum mechanical processes poses a problem to a regularity account of causation in the end boils down to what pre-theoretic understanding of the notion of a cause-effect relation is presupposed. Or put differently: Irreducibly probabilistic processes pose a problem to a regularity account only if such processes are to be identified as *causal* processes. If, however, the notion of a causal dependency is taken to be essentially tied to such principles as the principle of determinism, quantum mechanical processes are not classified as causal to begin with. Against such a conceptual background, indeterminism, rather than compromising regularity accounts, raises questions as to how it can "be the case in an indeterminist world that some events are causally determined while others are not".[31] Indeed, in view of irreducibly indeterministic processes, the regularity theorist might well retreat to a more moderate position according to which several types of causal relations can be distinguished, one of which – prevalent in macroscopic contexts – being a deterministic relation. Then he could propagate his account as analysis of just that deterministic causal relation.[32]

11 Conclusion

Of the standard arguments against regularity theories all but one target over-simplified theoretical sketches, which by no means conform to modern regularity theoretic analyses. Present-day regularity theories successfully handle imperfect, empty and single-case regularities, adequately represent epiphenomenal structures and capture the asymmetry of the causal relation. In short, they do justice to the whole complexity of causal structures. And all this is accomplished with simple recourse to extensional standard logic. Appropriate minimalization strategies are at hand such that redundancies implicit in material conditionals – e.g. due to monotony – can efficaciously be precluded from a causal interpretation. The theoretical price for this technical straightforwardness is that regularity theories have to subscribe to the principle of determinism which may well not be universally valid in light of standard interpretations of quantum mechanics. However, in view of the overall lack of a theoretical framework that can successfully causally interpret irreducibly indeterministic processes, regularity theories can be seen as a promising and very intuitive alternative to popular theoretical frameworks as implemented by counterfactual or probabilistic analyses. Moreover, if it is ultimately found that there in fact are indeterministic causal processes on micro level, regularity theories can be propagated as analyses of deterministic macroscopic causal dependencies.

Notes

[1]Cf. section 10.

[2]Cf. Baumgartner (forthcoming).

[3]There are some regularity theoretic proposals that do not subscribe to this tenet, e.g. Mackie (1965), (cf. section 9).

[4]This focus on events does not straightforwardly cover cases of causally related absences or omissions. The problems posed by causal dependencies as between omitted vaccination and contracting influenza will be neglected in the present context. They are treated in Baumgartner (forthcoming), ch. 3. For interesting proposals on how to deal with causation among absences cf. also Collins, Hall, and Paul (2004).

[5]The notion of an event type can be broadly spelled out in terms of event classes, such that instantiation of a type by an event token is easily explicable with reference to the membership relation. For details cf. Baumgartner (forthcoming), ch. 2.

[6]Hume originally required temporal succession, not mere proximity, for causally related events (cf. Hume (1999 (1748)), p. 146). In accordance with the usual practice, causes and effects are here only required to be spatiotemporally proximate such as not to preclude the possibility of simultaneous or backward causation on a priori grounds (cf. section 6).

[7]Broad (1930), p. 306.

[8]Cf. Hume (1978 (1740)), p. 105.

[9]Cf. Brand and Swain (1970), p. 222.

[10]Not all critics of regularity accounts have taken note of these purely logical ways to minimalize sufficient conditions. For instance, in 1970 Brand and Swain still erroneously claimed that minimalizing sufficient conditions cannot be accomplished in non-causal and, thus, non-circular terms (cf. Brand and Swain (1970), p. 226).

[11]Cf. e.g. Armstrong (1983).

[12]"Inexistent" in this context is not to be read as "has occurred prior to a specific moment of investigation", but rather is to be understood in terms of "has not occurred in all past and will not occur in all future". As mentioned above, the problem of causally interpreted omissions or absences shall be bypassed here. We are thus concerned with causation among positive factors only. (Cf. footnote 4 above).

[13]Cf. e.g. Armstrong (1983), ch. 2.

[14]Cf. e.g. Suppes (1970) or Price (1992).

[15]Cf. e.g. Reichenbach (1956).

[16](2) is a mere tentative formal representation, for, as mentioned in section 2, propositional logic does not allow for adequately expressing the relational constraints implicit in causal regularities in the sense of (VI). Cf. Graßhoff and May (2001) or Baumgartner (forthcoming) for details on the first-order representation of these dependencies.

[17]This essentially corresponds to Mackie's (1974) famous analysis of causation in terms of so-called *INUS-conditions*. Mackie (1974) will not be given an in-depth review in the present context. This has been done in Baumgartner and Graßhoff (2004), ch. 5.

[18]Cf. Graßhoff and May (2001), pp. 97-99. Similar analyses of the direction of causation have been proposed in Sanford (1976) or Hausman (1998).

[19]Cf. e.g. Cartwright (1989), pp. 25-29.

[20]The graphical notation implemented here is properly introduced in Baumgartner and Graßhoff (2004), ch. 3.
[21]Cf. Mackie (1974), p. 83 et seq., Cartwright (1989), pp. 25-29. In Baumgartner and Graßhoff (2004), pp. 99-103, we have discussed the Manchester-Hooters in all detail along with a solution to this problem.
[22]For more details cf. Baumgartner and Graßhoff (2004) or Baumgartner (forthcoming).
[23]Cf. May (1999), pp. 67-68.
[24]Cf. Armstrong (1983), ch. 1.
[25]For details about how a regularity theory along the lines of (VIII) deals with preemption cf. Graßhoff and May (2001), pp. 104-105.
[26]Nonetheless there also are deterministic interpretations of quantum mechanics (cf. Albert (1992)).
[27]Cf. van Fraassen (1989).
[28]Cf. Reichenbach (1956).
[29]Cf. Spirtes, Glymour, and Scheines (2000 (1993)).
[30]Cf. Spirtes, Glymour, and Scheines (2000 (1993)), or Cartwright (1999).
[31]Cf. Xu (1997), p. 137. Accordingly, Xu (1997) proposes a deterministic analysis of causation that he claims to be compatible with indeterminism. Similarly: Belnap (2005).
[32]Confining oneself to such a moderate position seems advisable for determinists in view of arguments in favor of a causal interpretation of irreducibly indeterministic processes as e.g. presented in Mellor (1995), ch. 5.

References

Albert, David Z.: *Quantum Mechanics and Experience*, Cambridge: Harvard University Press 1992.

Armstrong, David M.: *What is a Law of Nature?* Cambridge: Cambridge University Press 1983.

Baumgartner, Michael and Graßhoff, Gerd: *Kausalität und kausales Schliessen. Eine Einführung mit interaktiven Übungen*, Bern: Bern Studies 2004.

Baumgartner, Michael: *Complex Causal Structures. Extensions of a Regularity Theory of Causation*, forthcoming.

Beebee, Helen: Does Anything Hold the Universe Together? *Synthese* 149 (2006), 509–533.

Belnap, Nuel: A Theory of Causation: *Causae Causantes* (Originating Causes) as Inus Conditions in Branching Space-Times, *British Journal for the Philosophy of Science* 56 (2005), 221–253.

Brand, Myles and Swain, Marshall: On the Analysis of Causation, *Synthese* 21 (1970), 222–227.

Broad, Charles D.: The Principles of Demonstrative Induction I-II, *Mind* 39 (1930), 302–317, 426–439.

Cartwright, Nancy: *Nature's Capacities and Their Measurement*, Oxford: Clarendon Press 1989.

——— Causal Diversity and the Markov Condition, *Synthese* 121 (1999), 3–27.

Collins, John, Hall, Ned, and Paul, L.A., editors: *Causation and Counterfactuals*, Cambridge: MIT Press 2004.

Fraassen, Bas van: The Charybdis of Realism. Epistemological Implications of Bell's Inequality, in: Cushing, James T. and McMullin, Ernan, editors: *Philosophical Consequences of Quantum Theory*, Notre Dame: University of Notre Dame Press 1989, 97–113.

Graßhoff, Gerd and May, Michael: Causal Regularities, in: Spohn, W., Ledwig, M., and Esfeld, M., editors: *Current Issues in Causation*, Paderborn: Mentis 2001, 85–114.

Hall, Ned: Two Concepts of Causation, in: Collins, J., Hall, N., and Paul, L.A., editors: *Counterfactuals and Causation*, Cambridge: MIT Press 2004, 225–276.

Hausman, Daniel: *Causal Asymmetries*, New York: Cambridge University Press 1998.

Hitchcock, Christopher: Probabilistic Causation, in: Zalta, Edward N., editor: *The Stanford Encyclopedia of Philosophy*, 2002 ⟨URL: `http://plato.stanford.edu/archives/fall2002/entries/causation-probabilistic/`⟩.

Hume, David; Selby-Bigge, L. A. and Nidditch, P. H., editors: *A Treatise of Human Nature*, Oxford: Clarendon Press 1978 (1740).

——— *An Enquiry Concerning Human Understanding*, Oxford: Oxford University Press 1999 (1748).

Kim, Jaegwon: Causation, Nomic Subsumption, and the Concept of Event, *Journal of Philosophy* 70 (1973), 217–236.

Mackie, John L.: *The Cement of the Universe. A Study of Causation*, Oxford: Clarendon Press 1974.

——— Causes and Conditions, in: Sosa, Ernest and Tooley, Michael, editors: *Causation*, Oxford: Oxford University Press 1993 (1965), 33–55.

May, Michael: *Kausales Schliessen. Eine Untersuchung über kausale Erklärungen und Theorienbildung*, Ph. D thesis, Universität Hamburg 1999.

Mellor, David Hugh: *The Facts of Causation*, London: Routledge 1995.

Price, Huw: Agency and Causal Asymmetry, *Mind* 101 (1992), 501–520.

Reichenbach, Hans: *The Direction of Time*, Berkeley: University of California Press 1956.

Sanford, David: The Direction of Causation and the Direction of Conditionship, *The Journal of Philosophy* 73 (1976), 193–207.

Spirtes, Peter, Glymour, Clark, and Scheines, Richard: *Causation, Prediction, and Search*, 2nd edition. Cambridge: MIT Press 2000 (1993).

Suppes, Patrick: *A Probabilistic Theory of Causality*, Amsterdam: North Holland 1970.

Xu, Ming: Causation in Branching Time (I): Transitions, Events and Causes, *Synthese* 112 (1997), 137–192.