Boolean difference-making: a modern regularity theory of causation

Michael Baumgartner[*] and Christoph Falk[†]

Abstract

A regularity theory of causation analyses type-level causation in terms of Boolean difference-making. The essential ingredient that helps this theoretical framework overcome the well-known problems of Hume's and Mill's classical regularity theoretic proposals is a *principle of non-redundancy*: only redundancy-free Boolean dependency structures track causation. The first part of this paper argues that the recent regularity theoretic literature has not consistently implemented this principle, for it disregarded two important types of redundancies: *componential* and *structural* redundancies. The second part then develops a new variant of a regularity theory that does justice to all types of redundancies and, thereby, provides the first all-inclusive notion of Boolean difference-making.

Keywords: Regularity theory of causation, difference-making, Boolean causal models, configurational comparative methods, Coincidence Analysis, Qualitative Comparative Analysis

1 Introduction

Theories of causation come in many variants. Some take causation on the type level, *viz.* general causation or causal relevance, to be the primary analysandum, others focus on causation on the token level, *viz.* singular or actual causation. Some take some form of difference-making to be the primary analysans, others draw on the causes' powers or dispositions, yet others aim to spell out causation in terms of energy transfer or other features of the mechanism connecting causes and effects (for introductions to the different frameworks see

[*]University of Bergen, Norway, michael.baumgartner@uib.no
[†]University of Geneva, Switzerland, chri.falk@gmail.com

Beebee, Hitchcock, and Menzies 2009). Depending on these analytical choices, resulting theories ascribe different properties to causation. According to some theories, causation is a deterministic dependence relation, while according to others it is not; some stipulate that the relata of causation are spatiotemporally located entities, whereas for others absences of such entities may also be causally related; some theories presuppose or entail that causation is transitive, others entail that it is non-transitive; some yield that causation is a property that is extrinsic to causes and effects, according to others it is an intrinsic property; etc.

Although many of the available theories of causation are incompatible, philosophical research over the past decades has not led to the disappearance of any of them—arguably because they all track various stubborn pre-theoretic intuitions (e.g. Lombrozo 2010). No single consistent theory can possibly capture all the properties we pre-theoretically ascribe to causation. This insight, in turn, has induced a growing popularity of causal pluralism, *viz.* the view that there exist multiple concepts or variants of causation (e.g. Psillos 2010). A corollary of this view is that theories of causation should not be benchmarked against the whole range of popular pre-theoretic intuitions but only against maximally large consistent proper subsets of them. The selection of these subsets, of course, can vary from theory to theory and there is no fact of the matter as to what is the true or even best selection. Rather, the selection of features of causation a particular theory aims to capture must be justified on the basis of pragmatic considerations concerning the theoretical or methodological purposes a particular theory is intended to serve.

Accordingly, without claiming to be presenting the only and ultimate truth about causation, this paper develops a modern regularity theory of causation by correcting important shortcomings of the most recent regularity theoretical proposals. The resulting theory's purpose is to provide a conceptual fundament for the currently spreading configurational comparative methods (CCMs) of causal data analysis: *Qualitative Comparative Analysis* (QCA) (Rihoux and Ragin 2009; Thiem 2014), *Necessary Condition Analysis* (NCA) (Dul 2016), and *Coincidence Analysis* (CNA) (Baumgartner 2009a, 2009b).[1] CCMs differ from other techniques as regression analytical methods (RAMs) (e.g. Gelman and Hill 2007) or Bayes-nets methods (BNMs) (e.g. Spirtes, Glymour, and Scheines 2000) in a number of respects (for a discussion of some of these differences see Thiem et al. 2016). Most importantly for our current purposes, while RAMs and BNMs search for causal dependencies among *variables* by exploiting their statistical (in-)dependencies, CCMs search for causal dependencies

[1]CCMs have been applied in hundreds of studies in disciplines as diverse as business and economics, political science, international relations, sociology, management and organization, public health and health policy, education, environmental science, evaluation, legal studies, or media and communication. An overview over the literature is provided on the website of the COMPASSS network: `www.compasss.org`.

among concrete *values of variables* by exploiting Boolean dependencies as "A=α_i is (non-redundantly) sufficient/necessary for B=β_i" (where α_i and β_i are concrete values of A and B). To this end, CCMs must be underwritten by a theory of causation that provides a link between Boolean dependency structures and causation. This is exactly the field of expertise of regularity theories.

Metaphysically, regularity theories are embedded in the tradition of Humean actualist anti-necessitarianism (Hume 1748, sect. 7),[2] according to which there is no causal oomph enforcing the regularities obtaining in the world we live in; rather, causation, possibility, and lawhood supervene on the actual distribution of matters of fact, which itself is a brute fact of our world. Causal laws are convenient summaries of the regularities that happen to emerge from that distribution. Correspondingly, being in accordance with those laws, that is, being empirically possible is a matter of existing (in an atemporal sense) in the actual world (from its beginning to its end).

Analytically, regularity theories make the following choices. Their primary analysandum is causation on the type level, more specifically, causal relevance relations between variables or factors taking on specific values: "A=α_i is causally relevant to B=β_i", where A=α_i, for instance, stands for malfunctioning traffic lights and B=β_i for rear-end collisions.[3] (We will use the terms "variable" and "factor" interchangeably in this paper.) They adopt a classical difference-making approach that builds on the intuition that for each cause there must exist at least one fixed setting of context factors $\mathcal{F}$ in which it makes a difference to the effect, meaning that, in $\mathcal{F}$, a change in the cause is systematically associated with a change in the effect. In consequence, a causal relation between A=α_i and B=β_i is not taken to supervene on intrinsic properties of the (sets of) entities represented by A=α_i and B=β_i, rather it obtains in virtue of the factors' location and function in a whole network of other relevant factors. Moreover, causation is assumed to be a deterministic dependence relation, that is, the indeterminism in ordinary data is due our epistemic limitations and our resulting inability to control for all background noise. The primary analysans of regularity theories, then, consists in structures of Boolean dependencies of sufficiency and necessity that are rigorously freed of redundancies (Graßhoff and May 2001; Baumgartner 2008, 2013).

[2]As all metaphysical frameworks, actualist anti-necessitarianism, of course, is controversial. This paper, however, is not the place to enter that controversy. For its pragmatic scope, adopting the framework is sufficiently justified if that allows for developing a regularity theory that conceptually underwrites CCMs.

[3]There are some regularity theoretic proposals that consider token-level causation to be primary (e.g. Mackie 1965), but the criticism raised against these accounts (e.g. Kim 1971), in our view, shows that these accounts are beyond repair.

The principle, originally due to Broad (1930) and famously shaped in Mackie's (1974) INUS-theory, that only *redundancy-free* Boolean dependencies track causation, is the essential theoretical ingredient that helped overcome the well-known problems incurred by the classical regularity theoretic proposals (e.g. Hume 1748 and Mill 1843). To render this non-redundancy principle formally precise, Graßhoff and May (2001) determined that only Boolean dependencies in the form of *minimal biconditionals* track causation, where a minimal biconditional $\Pi \leftrightarrow B{=}\beta_i$ features an outcome $B{=}\beta_i$ on one side and a *minimally necessary* disjunction (in disjunctive normal form) of *minimally sufficient* conditions of $B{=}\beta_i$ on the other. Baumgartner (2013) suggested that this idea could be generalised for the analysis of structures with multiple outcomes by simply conjunctively concatenating minimal biconditionals. It was an (implicit) assumption of all these proposals that the only elements of Boolean dependency structures that can feature redundancies are sufficient and necessary conditions and that, accordingly, all that is required to do justice to the non-redundancy principle is to eliminate redundancies from relationships of sufficiency and necessity.

The first part of this paper, however, will show that this assumption is false. It is not true that minimizing sufficient and necessary conditions ensures the redundancy-freeness of Boolean dependency structures. There are two additional types of redundancies the regularity theoretic literature so far has disregarded: componential and structural redundancies. An outcome $B{=}\beta_i$ may have multiple minimal biconditionals, say, $\Pi_1 \leftrightarrow B{=}\beta_i$ and $\Pi_2 \leftrightarrow B{=}\beta_i$ such that Π_1 accounts for $B{=}\beta_i$ in terms of a proper subset of the factor values in Π_2. We shall say that $\Pi_2 \leftrightarrow B{=}\beta_i$ contains a *componential redundancy*, and we will show that a causal interpretation of minimal biconditionals with componential redundancies is not warranted because for componentially redundant factor values there cannot exist a fixed setting of context factors $\mathcal{F}$ in which they make a difference to the outcome. In return, the factor values contained in minimal biconditionals without componential redundancies are exactly those factor values for which difference-making contexts are possible, meaning that these biconditionals—which we shall label *RDN-biconditionals*—indeed are redundancy-free.

However, when RDN-biconditionals for single outcomes are conjunctively combined to multi-outcome RDN-biconditionals yet another form of redundancy may arise. An RDN-biconditional Φ_1, while internally free of redundancies, may be redundant in a conjunctive sequence $\Phi_1 * \Phi_2 * \ldots * \Phi_n$ of RDN-biconditionals because $\Phi_1 * \Phi_2 * \ldots * \Phi_n$ is logically equivalent to a proper part of itself: $\Phi_2 * \ldots * \Phi_n$. As Φ_1, hence, makes no difference to the overall structure, we shall say that Φ_1 is *structurally redundant*. In sum, what counts as a redundancy-free Boolean dependency structure that tracks causation does not only depend on the minimality of sufficient and necessary conditions but also on componential redundancy-freeness and on the redundancy-freeness of the conjunctive concatenation constituting the structure as a whole.

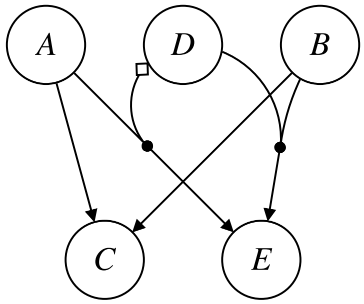
The second part of the paper then develops a new regularity theory that does justice to all types of redundancies and, thereby, provides the first all-inclusive notion of Boolean difference-making. Appendix 2 provides an R script that allows for easily replicating all analytical and calculative steps undertaken in this paper.

2 Background

A regularity theory assumes that its analysandum, *viz.* type-level causation, is not a fundamental property but that it supervenes on actual distributions of matters of fact, that is, on *Humean mosaics* (e.g. Lewis 1986, xi-x), which amount to *sets of configurations* of natural properties coincidently instantiated by units of observation—events, states of affairs, situations, cases, or whatever other entities the preferred ontology happens to furnish. The problem of rendering the notion of a natural property precise is notoriously difficult. For the purposes of this paper, we bracket that problem and simply assume that all henceforth analysed properties are natural. Moreover, as is common in the causal modelling literature, we want to remain as non-committal as possible with respect to the ontology of causation. We thus refer to the causal relata simply as "factors taking values".

Factors represent categorical properties that partition sets of units of observation either into two sets, in case of binary properties, or into more than two (but finitely many) sets, in case of multi-value properties. In the context of CCMs (e.g. Thiem 2014), factors representing binary properties are referred to as *crisp-set* or *fuzzy-set factors*; the former can take on the Boolean identity elements 0 and 1 as possible values, whereas the latter can take on any (continuous) values from the unit interval $[0, 1]$. Factors representing multi-value properties are called *multi-value factors*; they can take on any of an open (but finite) number of possible values $\{0, 1, 2, \ldots, n\}$. For simplicity of exposition, we confine ourselves to crisp-set factors in the context of this paper.

The focus on the crisp-set case allows us, for instance, to conveniently abbreviate the explicit "Variable=value" notation, which generates convoluted syntactic expressions with increasing model complexity. As is conventional in Boolean algebra, we shall write "A" for A=1 and "a" for A=0. While this shorthand significantly simplifies the syntax of causal models, it introduces a risk of misinterpretation, for it yields that the factor A and its taking on the value 1 are both expressed by "A". Disambiguation must hence be facilitated by the concrete context in which "A" appears. Accordingly, whenever we do not explicitly



(a)

#	A	B	C	D	E
σ_1	0	0	0	0	0
σ_2	0	1	1	0	0
σ_3	0	0	0	1	0
σ_4	1	0	1	1	0
σ_5	1	0	1	0	1
σ_6	1	1	1	0	1
σ_7	0	1	1	1	1
σ_8	1	1	1	1	1

(b)

Figure/Table 1: An exemplary causal structure (a) (where "•" symbolises *conjunction* and "$\diamond$" expresses *negation*) with a corresponding complete Humean mosaic (b).

characterise italicized Roman letters as "factors", we use them in terms of the shorthand notation. Moreover, we shall write "$A{*}B$" for the conjunction of A and B, "$A + B$" for the disjunction of A and B, "$A \to B$" for the conditional "If A then B" ($a + B$), and "$A \leftrightarrow B$" for the biconditional "A iff B" ($A{*}B + a{*}b$).

To have a concrete context for our ensuing discussion of the analytical tools needed by a regularity theory, consider the causal structure over the set of crisp-set factors $\mathbf{F}_1 = \{A, B, C, D, E\}$ in Figure 1a, which has two non-standard graphical elements that require introduction: arrows merged by "•" symbolise conjunctive relevance, and "$\diamond$" expresses that the negation of the factor at the tail of the arrow is relevant. That is, Figure 1a depicts a causal structure such that A and B are two alternative causes of C and $A{*}d$ and $B{*}D$ are two alternative causes of E. Factors C and E can, for instance, be thought of as representing the on/off states of two lights, while factors A, B, and D represent the on/off states of three switches. Light C is on iff at least one of the switches A or B is on, and light E is on iff A is on and D is off or B and D are both on.

We assume that the structure is deterministic and, for simplicity, that there are no hidden/unmeasured causal paths to C and E (meaning that the causal paths through A, B, and D are the only paths leading to C and E). It then follows that the elements of $\mathbf{F}_1$ can be instantiated in exactly the 8 types of configurations σ_1 to σ_8 listed in Table 1b. Type σ_1, for instance, represents a configuration where all factors take the value 0, type σ_2 a configuration where all but B and C take value 0, etc. Most logically possible configurations of the factors in $\mathbf{F}_1$ are determined to be inexistent by the structure in 1a. For example, A and B cannot be combined with c, for they causally determine C. Overall, if the behaviour of the factors in $\mathbf{F}_1$ is underwritten by the structure in Figure 1a, Table 1b lists all and only their empirically possible configurations. As there are no hidden causal paths, Table 1b contains a complete distribution of possible matters of fact. We shall say that Table 1b is the *complete Humean mosaic* (i.e. the complete set of configurations) for the structure in Figure 1a. By the lights of a regularity theory, that the causal dependencies in Figure 1a obtain means nothing over and above Table 1b being a complete Humean mosaic.

A regularity theory defines causation in terms of sufficiency and necessity relations among factors representing different natural properties that are *modally independent*, meaning logically and conceptually independent and not related in terms of metaphysical dependencies such as supervenience, constitution, grounding, etc. Subject to the nature of the involved factors, sufficiency and necessity relations can be given a classical or a fuzzy-logic rendering (cf. Baumgartner and Ambühl ming). In the context of this paper, we can confine ourselves to the classical rendering in terms of material implication: A is sufficient for B iff $A \rightarrow B$, and A is necessary for B iff $B \rightarrow A$. Plainly, most of these Boolean dependencies have nothing to do with causation. To illustrate, the configuration $A{*}B{*}D{*}E$ is sufficient for C in Table 1b, for this table does not feature the combination of $A{*}B{*}D{*}E$ and c. The same holds for $a{*}B{*}D{*}E$, $A{*}B{*}d{*}E$, etc. Moreover, the disjunctive concatenation of all sufficient conditions of C is necessary for C; that is, the following relations of sufficiency and necessity obtain among C and the other factors in $\mathbf{F}_1$:

$$A{*}B{*}D{*}E + a{*}B{*}D{*}E + A{*}B{*}d{*}E + A{*}b{*}d{*}E + A{*}b{*}D{*}e + a{*}B{*}d{*}e \;\leftrightarrow\; C \qquad (1)$$

(1) obviously does not track causation, as the factor E, for example, is part of every sufficient condition of C, but neither E nor e is causally relevant for C in the structure in Figure 1a. Still, some relations of sufficiency and necessity in fact reflect underlying causal dependencies: in Table 1b, for example, A and B are individually sufficient and their disjunction is necessary for C and they are the two alternative causes of E. Accordingly, the crucial problem to be solved by a regularity theory is to filter out those Boolean dependencies that track causation.

The main reason why most structures of Boolean dependencies do not reflect causation is that they tend to contain different types of redundancies—redundancies in sufficiency and necessity relations but also, as we shall see, componential and structural redundancies—, whereas structures of causal dependencies do not feature redundant elements. Every part of a causal structure makes a difference to the behaviour of that structure in at least one context. Accordingly, the regularity theoretic analysans must be required to be redundancy-free.

Non-redundancy (NR). Only Boolean dependency structures that do not contain any redundant elements track causation.

When applied to sufficient and necessary conditions, (NR) entails that whatever can be removed from such conditions without affecting their sufficiency and necessity is not a difference-maker and, hence, not causally relevant. Only *minimally* sufficient and *minimally* necessary conditions possibly track causation (Graßhoff and May 2001).

Minimal sufficiency. Let Σ be a conjunction of factor values $Z_1{*}\ldots{*}Z_n$ with $1 \leq n$. Σ is a *minimally sufficient condition* of B, iff

(a) The factors in Σ and B represent different natural and modally independent properties,

(b) $\Sigma \rightarrow B$, and

(c) for no proper part Σ' of Σ: $\Sigma' \rightarrow B$ (where a proper part of a conjunction is that conjunction reduced by at least one conjunct).

Minimal necessity. Let Π be a disjunction (in disjunctive normal form) of factor values $Z_1{*}\ldots{*}Z_g + \ldots + Z_m{*}\ldots{*}Z_n$ with $1 \leq n$. Π is a *minimally necessary condition* of B iff

(a) The factors in Π and B represent different natural and modally independent properties,

(b) $B \rightarrow \Pi$, and

(c) for no proper part Π' of Π: $B \rightarrow \Pi'$ (where a proper part of a disjunction is that disjunction reduced by at least on disjunct).

To illustrate, the first disjunct of (1), $A{*}B{*}D{*}E$, is not a minimally sufficient condition of C because it contains sufficient proper parts, for instance, $B{*}D{*}E$ is itself sufficient for C in Table 1b. But $B{*}D{*}E$ is likewise not minimally sufficient, as it also contains sufficient proper parts. Overall, C has three minimally sufficient conditions in Table 1b: A, B, and E.[4] Their disjunction is necessary for C, that is, $C \rightarrow A + B + E$. That necessary condition, however, still contains the spurious dependency between E and C.[5] The reason is that it does not amount to a minimally necessary condition, as it contains a necessary proper part, *viz.* $A + B$. Whenever factor C takes the value 1 in Table 1b, so do the factors A or B. The same does not hold for $B + E$ nor for $A + E$. Or differently, E is redundant to account for C because, whenever E is given, so is $A + B$. But the reverse does not hold: in configurations σ_2 and σ_4, $A + B$ is given but E is not.

[4] All calculations can be replicated using the R script available in Appendix 2.

[5] E is (at least) an INUS condition of C as defined by Mackie (1974, 62), whose INUS-theory is therefore forced to interpret E as a cause of C. This is an instance of the so-called Manchester Factory Hooters problem that ultimately induced Mackie to abandon the INUS-theory.

Minimally sufficient and necessary conditions are expressed in *minimal biconditionals*:[6]

Minimal biconditional. A biconditional $\Pi \leftrightarrow B$ is a minimal biconditional for B iff Π is a minimally necessary disjunction, in disjunctive normal form, of minimally sufficient conditions of B. (Π is the antecedent and B the consequent of the minimal biconditional.)

The following is a minimal biconditional for C entailed by Table 1b:

$$A + B \leftrightarrow C \tag{2}$$

(2) is the only minimal biconditional for C entailed by Table 1b and it correctly identifies the causes of C in Figure 1a. However, removing redundancies from sufficient and necessary conditions alone, albeit sufficient to pinpoint the causes of C, does not generally suffice to do justice to (NR). Boolean dependency structures can feature two further types of redundancies. We discuss the first of them in the next section and the second in section 4.

3 Componential redundancies

Zhang and Zhang (ms) have recently shown that a factor Z contained in the antecedent of a minimal biconditional for B is not guaranteed to be a difference-maker of B because factor Z may be functionally redundant to account for factor B. We agree, but contend that in order to ensure that minimal biconditionals track causation (as defined by regularity theories) it is not sufficient to only eliminate functionally redundant factors, rather, *componentially* redundant factor *values* need to be discarded. What a regularity theory is ultimately concerned about are not difference-making relations between factors but between specific values of factors.

To introduce the problem of componential redundancies, reconsider the set of configurations in Table 1b. That table does not only entail one but four minimal biconditionals for E:

$$A{*}d + B{*}D \leftrightarrow E \tag{3}$$

$$A{*}B + A{*}d + a{*}C{*}D \leftrightarrow E \tag{4}$$

$$A{*}B + B{*}D + b{*}C{*}d \leftrightarrow E \tag{5}$$

$$A{*}B + a{*}C{*}D + b{*}C{*}d \leftrightarrow E \tag{6}$$

[6] In the literature (e.g. Graßhoff and May 2001; Baumgartner 2013), minimal biconditionals are typically referred to as *minimal theories*. We prefer to reserve that label for biconditionals that do justice to (NR), which minimal biconditionals do not (see section 7 below).

(3) correctly reflects the causes of E in Figure 1a, whereas (4) to (6) do not. Prima facie, this might seem to be a case of mere empirical underdetermination resulting in a model ambiguity (cf. section 6 below). However, on closer inspection, (4) to (6) turn out to contain redundancies that (3) is free of. (3) accounts for E in terms of the four factor values in the set $\mathbf{M}_{(3)} = \{A, B, D, d\}$, while (4) to (6) account for E in terms *proper supersets* of $\mathbf{M}_{(3)}$: $\mathbf{M}_{(4)} = \{A, a, B, C, D, d\}$, $\mathbf{M}_{(5)} = \{A, B, b, C, D, d\}$, $\mathbf{M}_{(6)} = \{A, a, B, b, C, D, d\}$. When causally interpreted (3) identifies the elements of $\mathbf{M}_{(3)}$ to be causally relevant for E, whereas (4) to (6) claim that the elements $\mathbf{M}_{(3)}$ *plus* additional factor values are causally relevant for E. Table 1b, however, does not comprise difference-making contexts underwriting these additional relevancies. To make this precise, we introduce the notion of a *difference-making pair*:[7]

Difference-making pair. Let $\Pi \leftrightarrow B$ be a minimal biconditional entailed by a set of configurations δ, let A (resp. a) be contained in Π, and let $\mathbf{G}$ be the set of all factors featured in Π. A difference-making pair for the relevance of A (resp. a) for B is a pair of configurations $\{\sigma_i, \sigma_j\}$ in δ such that A (resp. a) and B are given in σ_i and a (resp. A) and b are given in σ_j and all factors in $\mathbf{G}\backslash\{A\}$ are constant in both σ_i and σ_j.

A difference-making pair for the relevance of A constitutes a context in which A is indispensable to account for a change in the outcome. Table 1b contains difference-making pairs for the relevance of all elements of $\mathbf{M}_{(3)}$. For instance, all factors in (3) except A and E are constant in $\{\sigma_2, \sigma_6\}$ while A and E both take the value 0 in σ_2 and the value 1 in σ_6. Or $\{\sigma_4, \sigma_5\}$ is a difference-making pair for the relevance of d, as D changes from 1 to 0 and E from 0 to 1 while all other factors in (3) are constant. By contrast, Table 1b does not contain difference-making pairs for the relevance of any of the additional elements of $\mathbf{M}_{(4)}$ to $\mathbf{M}_{(6)}$.

As we prove in Appendix 1, the existence of difference-making pairs and the sets of factor values accounting for outcomes in minimal biconditionals are very tightly connected.

Theorem 1. *Let* $\Pi \leftrightarrow B$ *be a minimal biconditional entailed by a set of configurations* δ*. Every factor value in* Π *has a difference-making pair in* δ *iff there does not exist a minimal biconditional* $\Pi' \leftrightarrow B$ *entailed by* δ *such that the factor values in* Π' *are a proper subset of the factor values in* Π.

That is, the reason why not all factor values contained in (4) to (6) have difference-making pairs in Table 1b is that Table 1b entails a minimal biconditional, *viz.* (3), that accounts for

[7] That notion is related to (and inspired by) Zhang and Zhang's (ms) notion of a *witnessing pair*, whose purpose however it is to underwrite the functional non-redundancy of factors and not—as the notion of a difference-making pair—the componential non-redundancy of concrete values of factors.

E in terms of a proper subset of the components of (4) to (6). The latter, hence, contain redundancies, which we refer to as *componential redundancies*. Minimal biconditionals without componential redundancies are componentially minimal:

Componential minimality. A minimal biconditional $\Pi \leftrightarrow B$ entailed by a set of configurations δ is componentially minimal relative to δ iff there does not exist a minimal biconditional $\Pi' \leftrightarrow B$ entailed by δ such that the factor values in Π' are a proper subset of the factor values in Π.

It follows from Theorem 1 that every factor value in the antecedent of a minimal biconditional that is componentially minimal has a corresponding difference-making pair. Having a difference-making pair is necessary for being causally relevant (although not, as section 4 will show, sufficient).

Minimal biconditionals that are free of componential redundancies are internally redundancy-free. To facilitate our ensuing discussion we furnish them with a label: *RDN-biconditionals* (redundancy-free disjunctive normal form biconditionals).

RDN-biconditional. A biconditional $\Pi \leftrightarrow B$ entailed by a set of configurations δ is an RDN-biconditional for B iff $\Pi \leftrightarrow B$ is a minimal biconditional that is componentially minimal relative to δ. (Π is the antecedent and B the consequent of the RDN-biconditional).

Table 1b entails exactly two RDN-biconditionals: (2) and (3). The behaviour of the other factors in the set $\mathbf{F}_1$ from Figure 1a cannot be expressed in terms of an RDN-biconditional, as neither A, B, nor D have minimally necessary conditions in $\mathbf{F}_1$. To see this, compare, for example, configurations σ_7 and σ_8 in Table 1b. In both of these configurations, the factors in $\mathbf{F}_1 \backslash \{A\}$ take constant values, but the value of A changes. Hence, A cannot be expressed as a function of $\mathbf{F}_1 \backslash \{A\}$. The same holds for B and D. In sum, Table 1b entails the conjunction of RDN-biconditionals in expression (7), and the value assignments to the factors in $\mathbf{F}_1$ that render (7) true are exactly the configurations listed in Table 1b.

$$(A + B \leftrightarrow C) * (A{*}d + B{*}D \leftrightarrow E) \tag{7}$$

The causal interpretation of RDN-biconditionals is straightforward: their antecedents are to be interpreted in terms of causes of their consequents; conjunctions stand for complex causes, disjunctions for alternative causal paths. Hence, (7) amounts to a correct Boolean representation of the common cause structure in Figure 1a. However, as the next section will show, minimising sufficiency and necessity relations and eliminating componential redundancies does not yet suffice to successfully capture all causal structures in regularity theoretic terms.

4 Structural redundancies

Apart from componential redundancies, Boolean dependency structures can feature yet another type of redundancy that the regularity theoretic literature has disregarded so far. RDN-biconditionals, although internally redundancy free, can themselves—as a whole—be redundant in superordinate structures and, hence, fail to track causation due to a higher-order violation of (NR). We shall speak of *structural redundancies* in that context. A case in point is the structure in Figure 2a, again over the set of crisp-set factors $\mathbf{F}_1 = \{A, B, C, D, E\}$, which is assumed to feature all causal paths leading to D and E. Table 2b is the complete Humean mosaic corresponding to that structure.

That mosaic not only entails the two RDN-biconditionals conforming to the causal structure in Figure 2a, *viz.* (8) and (9), but also an RDN-biconditional for C, *viz.* (10):

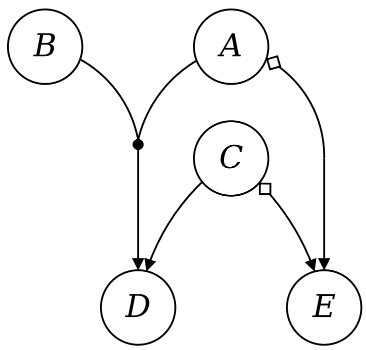
$$A{*}B + C \leftrightarrow D \tag{8}$$

$$a + c \leftrightarrow E \tag{9}$$

$$a{*}D + e \leftrightarrow C \tag{10}$$

That is, the behaviour of the factor C, which is exogenous in Figure 2a, can be expressed as an internally redundancy-free Boolean function of the two endogenous factors D and E. (10), hence, expresses a backtracking (or upstream) dependency. A regularity theory that only requires minimal sufficiency/necessity and componential minimality is forced to causally interpret that backtracking dependency, which, in turn, entails that the theory cannot always distinguish between causes and effects, thus vindicating a standard objection against regularity theories.

That consequence can be avoided by not only applying (NR) to sufficient and necessary conditions and their componential makeup but also to RDN-biconditionals as a whole. The backtracking dependency (10) is structurally redundant in the superordinate dependency structure (11) that results from a conjunctive concatenation of all the RDN-biconditionals that follow from Table 2b. The reason is that (11) has a proper part which is logically equivalent to (11), *viz.* (12).



(a)

#	A	B	C	D	E
σ_1	1	0	1	1	0
σ_2	1	1	1	1	0
σ_3	0	0	0	0	1
σ_4	1	0	0	0	1
σ_5	0	1	0	0	1
σ_6	1	1	0	1	1
σ_7	0	0	1	1	1
σ_8	0	1	1	1	1

(b)

#	A	C	D	E
σ_1	1	1	1	0
σ_2	1	1	1	0
σ_3	0	0	0	1
σ_4	1	0	0	1
σ_5	0	0	0	1
σ_6	1	0	1	1
σ_7	0	1	1	1
σ_8	0	1	1	1

(c)

Figure/Table 2: A complex causal structure (a) with a corresponding complete Humean mosaic (b) and an incomplete one (c).

$$(A{*}B + C \leftrightarrow D){*}(a + c \leftrightarrow E){*}(a{*}D + e \leftrightarrow C) \tag{11}$$

$$(A{*}B + C \leftrightarrow D){*}(a + c \leftrightarrow E) \tag{12}$$

Conjunctively adding (10) to (12) states nothing over and above what is already stated by (12), for (12) logically entails (10). (12) generates exactly the Humean mosaic in Table 2b, meaning that (10) is not required to account for that mosaic. At the same time, neither (8) nor (9) are redundant in (11), for neither the conjunction of (8) and (10) nor the conjunction of (9) and (10) is logically equivalent to (11)—neither of these conjunctions generates the configurations in Table 2b. Overall, the backtracking dependency (10) makes no difference to the generation of the Humean mosaic in Table 2b. It violates (NR); it is a structural redundancy in the context of the foretracking dependencies (8) and (9).

Thus, conjunctions of RDN-biconditionals only track causation if they are structurally minimal:

Structural minimality. A conjunction of RDN-biconditionals $\Psi = \Phi_1{*}\ldots{*}\Phi_n, 1 \leq n$, is structurally minimal iff there does not exist a Ψ' that results from Ψ by eliminating at least one conjunct such that Ψ and Ψ' are logically equivalent.

Every RDN-biconditional contained in a structurally minimal conjunction is *structurally indispensable*.

5 Permanence

Before assembling the analytical tools developed above in a new regularity theory, this section complements the minimality conditions by a permanence requirement and the next introduces an important metaphysical background assumption.

Real-life causal structures commonly are not as simple as the ones in Figures 1a and 2a. Single factors tend not to cause their effects in isolation. Rather, causes amount to very complex conjunctions of co-instantiated factors, which only jointly determine their effects. Moreover, on the type level, there typically exist more than two alternative paths to one effect. That is, regularities entailed by real-life structures tend to be much more complex than the ones expressed in (7) and (12). To adequately represent the complexity of real-life regularities, Mackie (1974, 66-71) uses X_1, X_2, etc. as placeholders for conjunctions of unmeasured (contextual) factor values and Y_1, Y_2, etc. as placeholders for disjunctions $X_1 + X_2 + \ldots + X_n$. Furthermore, to ensure that the complexity of Boolean dependency structures remains manageable, he (1974, 34-35, 63) relativizes regularities to what he calls a *causal field*, that is, to a fixed configuration of context factors. A more realistic scenario than the one in (12), thus, is that A, a, B, C, and c are mere parts of alternative causes of D and E within a field $\mathcal{F}$:

$$\text{in } \mathcal{F}: \; (A{*}B{*}X_1 \,+\, C{*}X_2 \,+\, Y_1 \;\leftrightarrow\; D) * (a{*}X_3 \,+\, c{*}X_4 \,+\, Y_2 \;\leftrightarrow\; E) \qquad (13)$$

In scientific discovery contexts, the constancy of the field, of course, is difficult to ensure, which is why real-life data will often not be as noise-free as Tables 1b and 2b. Hence, when causally analysing real-life data, strict Boolean dependencies can typically only be approximated. To this end, CCMs provide various parameters of model fit—the two most prominent ones being *consistency* and *coverage* (Ragin 2006). Roughly, consistency measures the degree to which a condition is sufficient for an effect, and coverage the degree to which a condition is necessary. By imposing consistency and coverage thresholds, it then becomes possible to treat conditions as sufficient or necessary even if, as in noisy data, Boolean dependencies are imperfect. However, since the focus of this paper is conceptual and not epistemological, we will not further discuss these issues here. Likewise, we abstain from making the field-relativity of regularities explicit and dispense with the explicit use of placeholders for open conjunctions and disjunctions. Instead, we do justice to the complexity of ordinary causal structures by assuming all Boolean dependency structures to be *open for expansions*, that is, for the integration of further factors.

The remainder of this section will show that expanding Boolean dependency structures provides an additional handle to constrain their causal interpretability—a handle that is needed because even Boolean dependency structures that satisfy all the minimality conditions introduced above may fail to actually track causation. The reason, in a nutshell, is that what counts as an RDN-biconditional is relative the analysed set of configurations and factors, and such sets may not faithfully reflect causation. For example, suppose that two factors A and B are instantiated in two configurations: in the first, A and B both take the value 1, and in the second, they both take the value 0. The corresponding set of configurations $\delta_1 = \{\langle 1,1\rangle, \langle 0,0\rangle\}$ entails the RDN-biconditional $A \leftrightarrow B$, which, if causally interpreted, suggests that A and B are causally related. If δ_1, however, does not contain all empirically possible configurations of A and B, $A \leftrightarrow B$ may fail to track causation. If, say, the set of all empirically possible configurations in fact is $\delta_2 = \{\langle 1,1\rangle, \langle 0,0\rangle, \langle 1,0\rangle, \langle 0,1\rangle\}$, A and B are causally independent, in which case a causal interpretation of $A \leftrightarrow B$ is false.

As anticipated in section 1, the anti-necessitarian tradition in which regularity theories are embedded provides an *actualist* rendering of the notion of an empirically possible configuration. That means the empirically possible configurations of the factors in a set $\mathbf{F}_\delta$ are all of their configurations that exist (in an atemporal sense) in the actual world (from its beginning to its end). Causation then supervenes on the actual distribution of matters of fact, which, in turn, is a brute fact of the world we live in. If a set of configurations δ contains all and only the empirically possible configurations of the factors in $\mathbf{F}_\delta$, we shall say that δ is an *exhaustive* set of configurations over $\mathbf{F}_\delta$. The above example thus shows that non-exhaustive sets may entail RDN-biconditionals that do not track causation. Still, when δ_1 is expanded to the exhaustive set δ_2, the spurious dependence between A and B disappears.

But exhaustiveness alone does not guarantee faithfulness to causation, as even structurally indispensable RDN-biconditionals entailed by exhaustive sets may fail to reflect causation. To see this, reconsider the structure in Figure 2a and assume that it is analysed without taking the factor B into account, that is, relative to $\mathbf{F}_2 = \{A, C, D, E\}$. Table 2c lists all empirically possible configurations of $\mathbf{F}_2$'s elements and, hence, amounts to an exhaustive set of configurations over $\mathbf{F}_2$. That table does not allow for expressing the behaviour of D as a function of $\mathbf{F}_2\backslash\{D\}$, because in the configurations σ_4 and σ_6 all factors in $\mathbf{F}_2\backslash\{D\}$ are constant while D changes. Table 2c only entails the RDN-biconditionals (9) and (10), whose conjunctive concatenation is

$$(a + c \leftrightarrow E) * (a * D + e \leftrightarrow C) \tag{14}$$

(14) does not contain a logically equivalent proper part, meaning that it structurally minimal and, accordingly, that its second conjunct (10) cannot be discarded based on $\mathbf{F}_2$.

Despite its structural minimality, (14) does not track causation, for C is not actually endogenous in the underlying structure of Figure 2a. The reason why (10) is not identified as a spurious upstream regularity is that $\mathbf{F}_2$ is *underspecified*, meaning that a causally relevant factor, which is not constant in the corresponding causal field, is left out of the analysis. In consequence, the Boolean dependencies among the elements of $\mathbf{F}_2$ cannot be completely freed of redundancies, even based on an exhaustive set of configurations.

Plainly, whether a factor set $\mathbf{F}_\delta$ is underspecified depends on what the causal structure is that underwrites the behaviour of $\mathbf{F}_\delta$'s elements. Accordingly, in the conceptual context of analysing causation or in the epistemic context of searching for the causal structure behind $\mathbf{F}_\delta$, $\mathbf{F}_\delta$ cannot be assumed to be free of underspecification (for this would presuppose clarity on causation and on the causal structure behind $\mathbf{F}_\delta$, respectively). Fortunately, neither the conceptual nor the epistemic context require such an assumption. The reason is that by gradually expanding factor sets spurious dependencies are identified in exhaustive sets of configurations. As soon as $\mathbf{F}_2$ is expanded to $\mathbf{F}_1$, D becomes expressible as a function of $\mathbf{F}_1 \backslash \{D\}$, meaning that (8) follows, which, as we have seen in the previous section, correctly reveals the redundancy of (10). Generally, dependencies that appear to be of the difference-making type relative to a set $\mathbf{F}_\delta$, but in fact are spurious, are identified as such in the course of gradual expansions of $\mathbf{F}_\delta$.

But in order to reliably reveal the spuriousness of Boolean dependencies, expansions of factor sets must be suitable for causal modelling. A *suitable expansion* $\mathbf{F}'_{\delta'}$ of a factor set $\mathbf{F}_\delta$ is a superset of $\mathbf{F}_\delta$, which is the result of introducing factors into $\mathbf{F}_\delta$ representing natural properties that are modally independent of one another and of the properties represented by the elements of $\mathbf{F}_\delta$. A suitable expansion $\mathbf{F}'_{\delta'}$ of $\mathbf{F}_\delta$ reveals that an RDN-biconditional $\Pi_i \leftrightarrow B$ over $\mathbf{F}_\delta$ features redundancies if there does not exist a structurally indispensable RDN-biconditional $\Pi_j \leftrightarrow B$ entailed by an exhaustive set of configurations δ' over $\mathbf{F}'_{\delta'}$ such that all components of Π_i are also components of Π_j. If there does not exist a suitable expansion $\mathbf{F}'_{\delta'}$ revealing redundancies in $\Pi_i \leftrightarrow B$, $\Pi_i \leftrightarrow B$ is *permanently redundancy-free*. A structurally indispensable RDN-biconditional tracks causation only if it is permanently redundancy-free.

#	A	B	C	D
σ_1	0	0	0	0
σ_2	0	0	1	0
σ_3	1	1	1	0
σ_4	1	0	0	1
σ_5	0	1	0	1
σ_6	1	1	0	1
σ_7	1	0	1	1
σ_8	0	1	1	1

(a)

#	A	B	C	D	E
σ_1	0	0	0	0	0
σ_2	1	0	0	0	0
σ_3	0	0	1	0	0
σ_4	1	0	1	0	0
σ_5	1	1	1	0	0
σ_6	0	1	0	1	0
σ_7	1	1	0	1	0
σ_8	0	1	1	1	0
σ_9	0	0	0	0	1
σ_{10}	0	0	1	0	1
σ_{11}	1	1	1	0	1
σ_{12}	1	0	0	1	1
σ_{13}	0	1	0	1	1
σ_{14}	1	1	0	1	1
σ_{15}	1	0	1	1	1
σ_{16}	0	1	1	1	1

(b)

$$A{*}b + a{*}B + A{*}c \leftrightarrow D \quad (15)$$
$$A{*}b + a{*}B + B{*}c \leftrightarrow D \quad (16)$$
$$A{*}b{*}E + a{*}B + B{*}c \leftrightarrow D \quad (17)$$

Table 3: (a) features a set of configurations that entails two RDN-biconditionals, *viz.* (15) and (16), whereas (b) results from (a) by expansion and entails only one RDN-biconditional, *viz.* (17).

6 Uniqueness

Apart from the metaphysical background assumptions concerning the nature of causation, most theories of causation additionally rely on auxiliary background assumptions in order to supply their analytical machineries with maximal traction. Just to mention one well-known example, an interventionist theory assumes that for all causally modelled variables there exist possible ideal interventions, which are surgical causes of those variables (Woodward 2003). A successful regularity theory also requires an auxiliary assumption.

To introduce it, consider the set of configurations in Table 3a over the factor set $\mathbf{F}_3 = \{A, B, C, D\}$. Table 3a entails two *logically equivalent* RDN-biconditionals for D, *viz.* (15) and (16), meaning that the behaviour of D cannot be unambiguously modelled on the basis of Table 3a. While both RDN-biconditionals—if causally interpreted—identify $A{*}b$ and $a{*}B$ as complex causes of D, it is indeterminate whether $A{*}c$ or $B{*}c$ are causally relevant to D. Plainly, ambiguities are a very widespread phenomenon in causal modelling within all modelling frameworks (e.g. Spirtes et al. 2000, 59-72; Baumgartner and Thiem 2015). Every so often, data do not unambiguously reflect causal structures. Such empirical underdetermination is normally taken to have epistemic sources, for instance, insufficient control over background factors or lacking surgicality of experimental manipulations. In principle, however, ambiguities can also originate from ideal data. If we assume Table 3a to be a complete Humean mosaic, that table is a case in point. Under this completeness assumption, Table 3a furnishes a perfectly noise-free and exhaustive set of configurations. That is, no further or cleaner empirical information could exist that would discriminate between (15) and (16), meaning there is no fact of the matter which of these RDN-biconditionals reflects the causal structure behind the behaviour of D. In other words, if causation is understood in regularity theoretic terms and Table 3a is a complete mosaic of a causally modelled world, that world does not have a determinate causal structure.

While it is easy to devise artificial toy worlds (e.g. in thought experiments targeting the adequacy of theories of causation) without determinate causal structures, we take it as a given that the world we live in is not of this kind. The causal makeup of our world may be beyond our epistemic reach, but it is ultimately *one* determinate makeup. Against the backdrop of this (metaphysical) uniqueness principle, Table 3a cannot amount to a complete mosaic. Rather, the ambiguity between (15) and (16) must be due to the fact that Table 3a does not contain all possible configurations of the factors in $\mathbf{F}_3$, but only all configurations *within some causal field* in which further relevant factors are constant. In fact, relative to an expanded factor set $\mathbf{F}_4 = \mathbf{F}_3 \cup E$ that results from $\mathbf{F}_3$ by integrating the factor E, the ambiguity can easily be resolved. To see this, consider Table 3b, which contains Table 3a as a proper part (highlighted with grey shading). Whenever the added factor E takes the value 1, the factors in $\mathbf{F}_3$ are instantiated in the configurations reported in Table 3a; but when E takes the value 0, further configurations are possible. Table 3b only entails one RDN-biconditional for D, *viz.* (17). That is, while it is impossible to determine whether D is caused by $A{*}c$ or $B{*}c$ relative to Table 3a, Table 3b resolves that ambiguity in favour of $B{*}c$.

As an auxiliary assumption we thus stipulate that model ambiguities are due to data deficiencies, rather than to an ultimate indeterminateness of the causal structure of the world we live in. In principle, thus, model ambiguities can always be resolved by expanding analysed factor sets. In other words, we assume causal uniqueness for complete Humean mosaics:

Causal Uniqueness (CU). Every complete Humean mosaic corresponds to one determinate causal structure.

7 A new regularity theory

We have now collected all ingredients for a new regularity theory of causation. To present that theory, we will proceed in two steps. First, we introduce the notions of a *minimal theory* and of an *atomic* minimal theory, and second, we define the notion of causal relevance in terms of membership in permanently redundancy-free atomic minimal theories.

Roughly, a minimal theory is a conjunction of (one or more) RDN-biconditionals that is structurally minimal. Subject to Causal Uniqueness (CU), every complete Humean mosaic, that is, every exhaustive set of configurations δ over a set of factors $\mathbf{F}_\delta$ without underspecification corresponds to exactly one causal structure Δ. Complete mosaics allow for complete redundancy elimination. Hence, a minimal theory entailed by a complete δ is free of all types of redundancies and, therefore, is guaranteed to truthfully reflect that Δ.

However, causal relevance cannot simply be defined in terms of minimal theories entailed by complete Humean mosaics. The reason was anticipated in section 5: clarity on the completeness of Humean mosaics presupposes clarity on the underlying causal structures, which is exactly what a theory of causal relevance is supposed to supply and thus, on pain of circularity, cannot presuppose. Accordingly, we shall not confine our notion of a minimal theory to complete mosaics. That, in turn, yields that minimal theories are not guaranteed to track causation. Nonetheless, the notion of a minimal theory shall be defined in such way that minimal theories amount to *representations of the causal evidence* contained in the set of configurations δ from which they follow. If δ is complete, that evidence is faithful to Δ, but if δ is incomplete, the evidence expressed by a minimal theory may misleadingly suggest the causal nature of some dependencies which in fact are spurious. Factor set expansions gradually rectify minimal theories entailed by a misleading δ by eliminating spurious dependencies and, thereby, 'zooming in' on the true Δ—thus the aforementioned second step in our analysis.

Here, then, is our definition of the notion of a minimal theory (simpliciter).

Minimal Theory. Let δ be a set of configurations over a factor set $\mathbf{F}_\delta$ and let $\Gamma = \Phi_1 * \ldots * \Phi_n$, $n \geq 1$, be the conjunction of all RDN-biconditionals entailed by δ. A minimal theory for δ over $\mathbf{F}_\delta$ is a conjunction $\Psi = \Phi_k * \ldots * \Phi_m, 1 \leq k \leq m \leq n$, of RDN-biconditionals from Γ such that the following conditions hold:

(a) Ψ is logically equivalent to Γ,

(b) Ψ is structurally minimal,

(c) any two Φ_i and Φ_j in Ψ have different consequents.

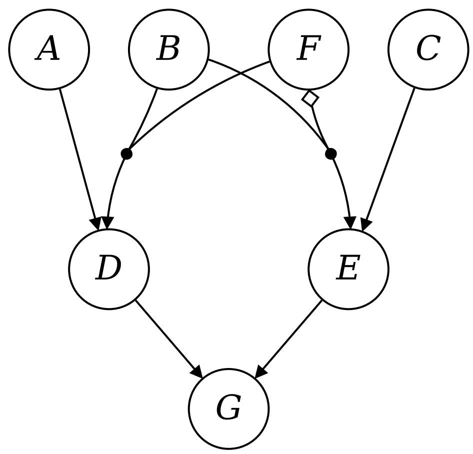
While the purpose of condition (b) is clear (see section 4), conditions (a) and (c) require explication. We begin with condition (a). A minimal theory shall express *all* the causal evidence contained in δ; Γ is the conjunction of *all* the RDN-biconditionals entailed by δ; the only RDN-biconditionals in Γ that do not express causal evidence are the structurally redundant ones filtered out in condition (b); but the elimination of structurally redundant RDN-biconditionals is an equivalence transformation; therefore, a minimal theory for δ must have the same truth conditions as Γ, meaning it must be logically equivalent to Γ.

To understand condition (c), recall from the previous section that sets of configurations sometimes entail multiple RDN-biconditionals with identical consequents. In order for a minimal theory Ψ to exhibit the causal evidence contained in a set δ, Ψ must be a candidate representation of the complete causal structure behind δ, that is, it must conjunctively concatenate as many RDN-biconditionals entailed by δ as can be interpreted as an integrated causal structure. To this end, it must be ensured that none of the RDN-biconditionals in Ψ have identical consequents, for RDN-biconditionals with an identical consequent B do not represent one causal structure but an ambiguity with respect to the causal structure behind B.[8] Condition (c) thus determines that a minimal theory contains maximally one RDN-biconditional for every endogenous factor. As a result, a Γ for a given δ may be broken down into multiple minimal theories. Table 3a is a case in point, as it entails two RDN-biconditionals, *viz.* (15) and (16), whose conjunctive concatenation does not amount to one minimal theory due to a violation of condition (c). Table 3a entails two different minimal theories: (15) and (16). The overall causal inference to be drawn from a δ entailing multiple minimal theories Ψ_1 to Ψ_n is *disjunctive*: the evidence in δ is such that Ψ_1 *or* Ψ_2 *or* ... *or* Ψ_n corresponds to the underlying causal structure Δ.[9]

A minimal theory entailed by a set δ comprehensively and rigorously implements the non-redundancy principle (NR) relative to δ, which, since Mackie (1974), counts as the core analytical tool on the way towards a successful regularity theory of causation. But as defined above, the tool is still too coarse-grained. To identify relations of Boolean difference-making we need the more fine-grained notion of an *atomic* minimal theory. To see why, consider the switching structure in Figure 4a with the corresponding exhaustive set of configurations in Table 4b. The ultimate effect of that structure, G, has two alternative causes, $D + E$, which

[8] Two RDN-biconditionals of B cannot be integrated into the same causal structure because each individually exhibits a minimally necessary disjunction of minimally sufficient conditions of B. Causally interpreting both at the same time would hence induce a redundancy, in violation of (NR).

[9] Note that the conjunction of (15) and (16) also violates condition (b), for (15) and (16) are logically equivalent. It does not hold generally, however, that RDN-biconditionals with identical consequents are logically equivalent.



(a)

#	A	B	C	D	E	F	G
σ_1	0	0	0	0	0	0	0
σ_2	0	0	0	0	0	1	0
σ_3	1	0	0	1	0	0	1
σ_4	0	1	0	0	1	0	1
σ_5	0	0	1	0	1	0	1
σ_6	0	1	1	0	1	0	1
σ_7	1	1	0	1	1	0	1
σ_8	1	0	1	1	1	0	1
σ_9	1	1	1	1	1	0	1
σ_{10}	1	0	0	1	0	1	1
σ_{11}	0	1	0	1	0	1	1
σ_{12}	1	1	0	1	0	1	1
σ_{13}	0	0	1	0	1	1	1
σ_{14}	1	0	1	1	1	1	1
σ_{15}	0	1	1	1	1	1	1
σ_{16}	1	1	1	1	1	1	1

(b)

Figure/Table 4: A switching structure with switch F (a) and corresponding exhaustive set of configurations (b).

themselves have two alternative causes each, $A + B{*}F$ for D and $C + B{*}f$ for E. The crucial feature of this structure is that factor F functions as a switch for the causal impact of B on D and E. The combination of B and F causes D, and the combination of B and f causes E. But independently of whether F is instantiated, B is sufficient to bring about G—via D in case of F and via E in case of f. Hence, the factor F only makes a difference to whether the causal influence of B on G is mediated by D or by E but not to G itself. Nonetheless, F and f appear in minimal theories representing causal structures with G as ultimate outcome. In total, Table 4b entails four minimal theories:

$$(A + B{*}F \leftrightarrow D) * (C + B{*}f \leftrightarrow E) * (D + E \leftrightarrow G) \tag{18}$$

$$(A + B{*}F \leftrightarrow D) * (C + B{*}f \leftrightarrow E) * (A + B + C \leftrightarrow G) \tag{19}$$

$$(A + B{*}F \leftrightarrow D) * (C + B{*}f \leftrightarrow E) * (A + B + E \leftrightarrow G) \tag{20}$$

$$(A + B{*}F \leftrightarrow D) * (C + B{*}f \leftrightarrow E) * (B + C + D \leftrightarrow G) \tag{21}$$

That the factor F is contained in these minimal theories apparently must not be taken to entail that F is a difference-maker of G. The reason is that, even though F is contained in minimally sufficient conditions of D and E, it is not part of a minimally sufficient condition of G—as B alone is sufficient for G. Correspondingly, F does not appear in an RDN-biconditional for G in any of the theories (18) to (21). That, in turn, shows that it is not membership in minimal theories (simpliciter) that tracks difference-making relations, but membership in RDN-biconditionals contained in minimal theories, which, for simplicity, we label *atomic* minimal theories:

Atomic Minimal Theory. An atomic minimal theory Φ of B entailed by a set of configurations δ over the set of factors $\mathbf{F}_\delta$ is an RDN-biconditional for B contained (as a conjunct) in a minimal theory Ψ for δ over $\mathbf{F}_\delta$.

To illustrate, the atomic minimal theories entailed by Table 4b are all the conjuncts of (18) to (21).

Against that background, we submit the following regularity theoretic definition of causal relevance (type-level causation):

Causal Relevance (CR). A is causally relevant for B iff there exists a set of (modally independent) factors $\mathbf{F}_\delta$, $A, B \in \mathbf{F}_\delta$, such that δ is the exhaustive set of configurations over $\mathbf{F}_\delta$, and the following conditions hold:

(a) A is part of the antecedent of an atomic minimal theory of B for δ over $\mathbf{F}_\delta$,

(b) for every suitable expansion $\mathbf{F'}_{\delta'} \supset \mathbf{F}_\delta$ and corresponding exhaustive set δ': A is contained in an atomic minimal theory of B for δ' over $\mathbf{F'}_{\delta'}$.

The core of (CR) can be less formally expressed as follows: A is causally relevant for B iff A is part of a permanently redundancy-free atomic minimal theory of B.

Before discussing the relevant implications of (CR) in the next section, let us highlight one important feature of (CR). Causal relevance as defined in (CR) is *non-transitive*. That is, it is possible for Z_1 to be causally relevant for Z_2, which itself is causally relevant for Z_3, without Z_1 being causally relevant for Z_3. The switching structure in Figure 4a is a case in point. Presuming that the minimal theory (18) is permanently redundancy-free, it follows that F is causally relevant for both D and E, which are relevant for G, but F is not causally relevant for G.

8 Discussion

While (CR) draws on analytical tools from previous regularity theoretic proposals (Mackie 1974, Graßhoff and May 2001; Baumgartner 2008, 2013), it assembles these tools in a way that implicates a departure from an implicit consensus underlying preceding regularity theories. All of its predecessors entail (or presuppose) that multi-outcome structures can be modularly built up from single-outcome structures, meaning that, in order to determine what the causes are of some effect B, it suffices to identify members of permanently redundancy-free sufficient and necessary conditions of B. In light of the arguments presented in this paper, that modularity principle can no longer be sustained. According to (CR), the redundancy-freeness of Boolean dependency structures and, thus, their causal interpretability cannot be assessed for single-outcome structures individually but only for complete multi-outcome structures taken together. As a result, (CR) entails a form of *causal holism* according to which causation is a holistic property that supervenes on complete Humean mosaics and not on proper parts thereof.

That holism has a number of notable ramifications. For instance, it yields that (CR) is more restrictive in sanctioning the causal interpretability of Boolean dependency structures than its predecessor theories: all dependencies that can be causally interpreted according to (CR) can also be causally interpreted according to its predecessors but not vice versa. This is particularly important in light of the fact that most of the classical objections levelled against regularity theories of type causation since the times of Hume and Mill contend that these theories overgenerate, meaning they stipulate the causal interpretability of Boolean dependencies which in fact are spurious. It follows that those overgeneration problems that have already been solved by (CR)'s predecessors are solved correspondingly by (CR); this concerns in particular the problems of empty and single-case regularities (cf. e.g. Armstrong 1983) as well as the notorious 'Manchester Factory Hooters' problem (Mackie 1974, 83-87). For detailed discussions of these issues the reader is hence referred to Graßhoff and May (2001) and Baumgartner (2008, 2013).

Sections 3 and 4 have shown, however, that two overgeneration problems have not been addressed by (CR)'s predecessors. First, Boolean dependency structures with only minimised sufficiency and necessity relations may contain factor values for which no difference-making contexts exist, that is, they may contain redundant components. In consequence, the theory developed in this paper introduces the notion of an RDN-biconditional which additionally requires Boolean dependency structures to be componentially minimal. Second, it can happen that the behaviour of factors, which are exogenous in the underlying causal structure, are expressible in terms of RDN-biconditionals featuring their own effects, which can be seen as a sophisticated version of the classical problem of distinguishing between upstream and downstream regularities. (CR) addresses this issue by prohibiting the causal interpretation of RDN-biconditionals that are structurally redundant, which eliminates upstream RDN-biconditionals.

Of course, regularity theories have also been objected to on the ground that they undergenerate in case of irreducible indeterminism, which does not result from our epistemic and practical limitations but is inherent in the underlying physical reality (cf. e.g. Dowe and Noordhof 2004). While standard interpretations of quantum mechanics advocate the existence of irreducibly indeterministic processes, non-standard interpretations disagree. Hence, there is no consensus on whether our universe is deterministic or not. Moreover, even if irreducibly indeterministic processes exist, there are many open questions—as for instance raised by phenomena of the EPR type—with respect to the *causal* interpretability of these processes (cf. e.g. Healey 2010). In the present context, we can sidestep these foundational questions, for, as indicated in the introduction, regularity theories aim to capture the intuition that causation is a deterministic form of dependence, that is, they analyse deterministic variants of causation (only). That means the notion of causal relevance spelled out in (CR) must be understood in terms of *deterministic* causal relevance. If there should turn out to exist irreducibly indeterministic causal relevance relations, other theoretical frameworks would have to be called upon.

Another upshot of the causal holism entailed by (CR) is that unmistakable inferences on causal relevance relations can only be drawn from complete (i.e. fully expanded) Humean mosaics, that is, from sets of configurations featuring factors on all paths involved in an analysed causal structure and comprising all empirically possible configurations of these factors. It is, of course, questionable whether complete Humean mosaics for other than artificial causal structures are ever available to human reasoners. Still, atomic minimal theories inferred from exhaustive sets of configurations amount to transparent representations of the causal evidence contained in those sets. Even though, in the absence of complete mosaics, causal inferences always run a risk of being refuted in the light of factor set expansions, such inferences become increasingly warranted the longer memberships in minimal theories are stable throughout a series of factor set expansions. That is, the inference to causal relevance as defined by (CR) is *inherently inductive*, which—we contend—nicely captures the nature of causal inference in scientific practice.

Finally, as (CR) is the first regularity theory that eliminates componential and structural redundancies and, thus, provides the first notion of Boolean difference-making that rigorously implements the non-redundancy principle, configurational comparative methods of causal data analysis, as QCA or CNA, which output Boolean dependency structures, are well-advised to understand causal relevance relations in terms of (CR). This, however, demands methodological adaptations from both of these methods. To date, neither QCA nor CNA eliminate componential redundancies from their Boolean causal models. Moreover, QCA focuses on single-outcome structures only and considers embedding single-outcome structures in superordinate multi-outcome structures as being optional. (CR) calls for a revision of that approach. Reliable Boolean causal inference not only requires expanding and improving the evidence base on the causes of single outcomes, but necessitates also aggregating single- to multi-outcome structures. While such an aggregation has always been an essential element in the procedural protocol of CNA, CNA has, so far, conceived of this aggregation in too simplistic a manner: it solely conjunctively concatenates minimal biconditionals inferred from processed data. According to (CR), however, an additional iteration of redundancy elimination is required: complex Boolean dependency structures must themselves be freed of redundant elements before they are amenable to a causal interpretation.

We end with three caveats. First, note that (CR) does not distinguish between *direct* and *indirect* causal relevance. In light of the non-transitivity of causal relevance as defined by (CR), indirect relevance cannot simply be spelled out in terms of the transitive closure of direct relevance, which, in turn, is accounted for in terms of membership in permanently redundancy-free atomic minimal theories. Discriminating between direct and indirect relevance presupposes a notion of a causal chain, which, for reasons of space, we cannot properly introduce here. Second, and on a related note, (CR) does not distinguish between *conjunctive* and *disjunctive* relevance relations, meaning it does not group causally relevant factor values into complex and alternative causes. To this end, a notion of a permanent minimal theory would be required, which—we reckon—should be easily obtainable by generalising the notion of a permanently redundancy-free minimal theory, but which we nonetheless have to leave for later. Third, note again that (CR) provides a notion of *type-level* causation. Token-level causation or actual causation must be cashed out in terms of a suitable spatiotemporal instantiation of a type-level structure. Building a corresponding token-level account on the basis of (CR) must also await another occasion.

Appendix 1

In this appendix, we prove Theorem 1, p. 10. To this end, let $\Pi \leftrightarrow B$ be a minimal biconditional entailed by a set of configurations δ. Let $\mathbf{G}$ be the set of factors, (at least) one of whose values is contained in Π, and let $\mathbf{M}$ be the set of those factor values in Π. Theorem 1 states the following equivalence:

(i) Every factor value in Π has a difference-making pair in δ.

$\leftrightarrow$

(ii) There does not exist a minimal biconditional $\Pi' \leftrightarrow B$ entailed by δ such that the factor values in Π' are a proper subset of the factor values in Π.

(i) $\rightarrow$ (ii):

(i) entails that for every factor value $Z \in \mathbf{M}$ there is a pair of configurations $\{\sigma_i, \sigma_j\}$ in δ such that Z and B are given in σ_i while z and b are given in σ_j, and all factors in $\mathbf{G}\backslash\{Z\}$ are constant in $\{\sigma_i, \sigma_j\}$. Now, assume for reductio that (ii) is false. It follows that δ entails a minimal biconditional $\Pi' \leftrightarrow B$ such that Π' contains the factor values $\mathbf{M}\backslash\{Z\}$ and such that Π' is true in σ_i and false in σ_j. As all factors in $\mathbf{G}\backslash\{Z\}$ are constant in $\{\sigma_i, \sigma_j\}$, Π' can only have different truth values in σ_i and σ_j if it contains the factor value z. But as Z is given in σ_i, every disjunct in Π' containing z is false in σ_i. Either Π' is true in σ_i nonetheless, because a disjunct not containing z is true in σ_i, or Π' as a whole is false in σ_i. In the first case, however, the truth of Π' is due to a disjunct containing only factors from $\mathbf{G}\backslash\{Z\}$, which—in light of the constancy of $\mathbf{G}\backslash\{Z\}$—will also render Π' true in σ_j. It follows that Π' cannot be true in σ_i and false in σ_j, which contradicts the assumption that (ii) is false. It follows that if (i) is true, (ii) must be true as well.

(ii) $\rightarrow$ (i):

We prove the contraposition $\neg$(i) $\rightarrow$ $\neg$(ii). If there does not exist a difference-making pair for at least one factor value Z in Π, it follows that there is no pair of configurations $\{\sigma_i, \sigma_j\}$ in δ such that Z and B are given in σ_i while z and b are given in σ_j, and all factors in $\mathbf{G}\backslash\{Z\}$ are constant in $\{\sigma_i, \sigma_j\}$. $\{\sigma_i, \sigma_j\}$ can fail to be such a difference-making pair in four different ways:

(1) $\mathbf{G}\backslash\{Z\}$ is not constant in $\{\sigma_i, \sigma_j\}$;

(2) Z is constant in $\{\sigma_i, \sigma_j\}$;

(3) B is constant in $\{\sigma_i, \sigma_j\}$;

(4) z and B are given in σ_i while Z and b are given in σ_j and $\mathbf{G}\backslash\{Z\}$ is constant in $\{\sigma_i, \sigma_j\}$.

(1) If $\mathbf{G}\backslash\{Z\}$ is not constant in $\{\sigma_i, \sigma_j\}$, any variation of B in $\{\sigma_i, \sigma_j\}$ can be accounted for by a corresponding variation in $\mathbf{G}\backslash\{Z\}$, meaning that factor Z is not needed to account for that variation of B. (2) If Z is constant in $\{\sigma_i, \sigma_j\}$, Z cannot account for a variation B in $\{\sigma_i, \sigma_j\}$, which accordingly must be accounted for by a variation in $\mathbf{G}\backslash\{Z\}$. (3) If B is constant in $\{\sigma_i, \sigma_j\}$, there is no variation of B to be accounted for, for which Z might be needed. (4) describes a difference-making pair for the relevance of z for B, meaning—as we

have seen above—that z is needed to account for the variation in B. In all cases (1) to (4), it holds that the factor value Z is not needed to account for a variation in B. As these four cases cover all possible types of pairs of configurations violating (i), the negation of (i) entails that the behavior of B can be expressed in terms of a biconditional not featuring the factor value Z. Since any such biconditional can be brought into a minimised disjunctive normal form, it follows that if (i) is false, there exists a minimal biconditional $\Pi' \leftrightarrow B$ entailed by δ such that the factor values in Π' are $\mathbf{M} \setminus \{Z\}$, meaning that the factor values in Π' are a proper subset of Π. Overall, if (i) is false, (ii) must be false as well, which, by contraposition, proves (ii) $\rightarrow$ (i).

Appendix 2

```
# R replication script
# ####################

# Required R package
library(cna)

# Table 1b
# --------
dat1 <- allCombs(c(2,2,2,2,2)) -1
(tab1b <- selectCases("(A + B <-> C)*(A*d + B*D <-> E)", dat1))
# minimally sufficient conditions for C
ana1 <- cna(tab1b, what="mac")
subset(msc(ana1), outcome=="C")
# one minimal biconditional for C
subset(asf(ana1), outcome=="C")
# four minimal biconditionals for E
subset(asf(ana1), outcome=="E")

# Table 2b
# --------
(tab2b <- selectCases("(A*B + C <-> D)*(c + a <-> E)", dat1))
# three RDN-biconditionals entailed by tab2b
ana2 <- cna(tab2b)
asf(ana2)
# conjunction of RDN-biconditionals with structural redundancy
csf(ana2)
# structurally minimal conjunction of RDN-biconditionals
minimalizeCsf(csf(ana2)$condition, dat1)
```

```
# Table 2c
# --------
(tab2c <- tt2df(tab2b[,-2]))
ana3 <- cna(tab2c)
# structurally minimal conjunction of RDN-biconditionals
minimalizeCsf(csf(ana3)$condition, dat1)

# Tables 3a/b
# -----------
dat2 <- allCombs(c(2,2,2,2)) -1
(tab3a <- selectCases("A*b + a*B + A*c <-> D", dat2))
# two structurally minimal RDN-biconditionals
cna(tab3a)
# ambiguity resolution through factor set expansion
dat3 <- allCombs(c(2,2,2,2,2)) -1
(tab3b <- selectCases("A*b*E + a*B + B*c <-> D", dat3))
cna(tab3b)

# Table 4b
# --------
dat4 <- allCombs(c(2,2,2,2,2,2,2)) -1
(tab4b <- selectCases("(A + B*F <-> D)*(C + B*f <-> E)*(D + E <-> G)",
                      dat4))
cna(tab4b)
```

References

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

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

Baumgartner, M. (2009a). Inferring causal complexity. *Sociological Methods & Research 38*, 71–101.

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

Baumgartner, M. (2013). A regularity theoretic approach to actual causation. *Erkenntnis 78*, 85–109.

Baumgartner, M. and M. Ambühl (forthcoming). Causal modeling with multi-value and fuzzy-set coincidence analysis. *Political Science Research and Methods*.

Baumgartner, M. and A. Thiem (2015). Model ambiguities in configurational comparative research. *Sociological Methods & Research*. doi: 10.1177/0049124115610351.

Beebee, H., C. Hitchcock, and P. Menzies (Eds.) (2009). *The Oxford Handbook of Causation*. Oxford: Oxford University Press.

Broad, C. D. (1930). The principles of demonstrative induction i-ii. *Mind 39*, 302–317, 426–439.

Dowe, P. and P. Noordhof (Eds.) (2004). *Cause and Chance. Causation in an Indeterministic World*, London. Routledge.

Dul, J. (2016). Necessary Condition Analysis (NCA): Logic and methodology of "necessary but not sufficient" causality. *Organizational Research Methods 19*(1), 10–52.

Gelman, A. and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

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

Healey, R. (2010). Causation in quantum mechanics. In H. Beebee, C. Hitchcock, and P. Menzies (Eds.), *The Oxford Handbook of Causation*. Oxford University Press. doi: 10.1093/oxfordhb/9780199279739.003.0034.

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

Kim, J. (1971). Causes and events: Mackie on causation. *Journal of Philosophy 68*, 426–441.

Lewis, D. (1986). *Philosophical Papers II*. Oxford: Oxford University Press.

Lombrozo, T. (2010). Causal–explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology 61*, 303–332.

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

Mackie, J. L. (1993 (1965)). Causes and conditions. In E. Sosa and M. Tooley (Eds.), *Causation*, pp. 33–55. Oxford: Oxford University Press.

Mill, J. S. (1843). *A System of Logic*. London: John W. Parker.

Psillos, S. (2010). Causal pluralism. In R. Vanderbeeken and B. D'Hooghe (Eds.), *World-views, Science and Us: Studies of Analytical Metaphysics*, pp. 131–151. World Scientific Publishers.

Ragin, C. C. (2006). Set relations in social research: Evaluating their consistency and coverage. *Political Analysis 14*, 291–310.

Rihoux, B. and C. C. Ragin (Eds.) (2009). *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks: Sage.

Spirtes, P., C. Glymour, and R. Scheines (2000). *Causation, Prediction, and Search* (2 ed.). Cambridge: MIT Press.

Thiem, A. (2014). Unifying configurational comparative methods: Generalized-set Qualitative Comparative Analysis. *Sociological Methods & Research 43*(2), 313–337.

Thiem, A., M. Baumgartner, and D. Bol (2016). Still lost in translation! A correction of three misunderstandings between configurational comparativists and regressional analysts. *Comparative Political Studies 49*, 742–774. doi: 10.1177/0010414014565892.

Woodward, J. (2003). *Making things happen. A theory of causal explanation*. Oxford: Oxford University Press.

Zhang, J. and K. Zhang (ms). Causal minimality in the Boolean approach to causal inference. manuscript.