Inferring Causal Complexity

Michael Baumgartner
University of Bern, Switzerland

In *The Comparative Method*, Ragin (1987) outlined a procedure of Boolean causal reasoning operating on pure coincidence data that has since become widely known as QCA (Qualitative Comparative Analysis) among social scientists. QCA—also in its recent forms as presented in Ragin (2000, 2008)—is designed to analyze causal structures featuring no more than *one* effect and a possibly complex configuration of *mutually independent direct* causes of that effect. The paper at hand presents a procedure of causal reasoning that operates on the same type of empirical data as QCA and that implements Boolean techniques related to the ones resorted to by QCA. Yet in contrast to QCA, the procedure introduced here successfully identifies structures involving both *multiple* effects and *mutually dependent* causes. In this sense, the paper at hand generalizes QCA.

KEY WORDS: causal discovery; data analysis; causal complexity; Boolean causal reasoning; Qualitative Comparative Analysis (QCA).

1. INTRODUCTION

In *The Comparative Method*, Ragin (1987) developed a methodology of causal analysis that has since become known as QCA (Qualitative Comparative Analysis) among social scientists. Ragin has introduced QCA as an alternative to standard quantitative and qualitative methodologies prevalent in social sciences. Social scientists are often confronted with data sets that are too small and too inhomogeneous for a significant statistical analyzability, or, as Mahoney and Goertz (2006) have argued, they are explicitly interested in sufficient or necessary causes rather than in statistical causal dependencies. At the same time, however, small-N data sets are

Author's Note: My particular thanks go to Timm Lampert and Daniel Engler, in collaboration with whom first attempts at adapting Quine-McCluskey optimization for the discovery of complex causal structures have been made. Moreover, I am grateful to Clark Glymour, Manuel Jakob, Tim Raez, Charles Ragin, and three anonymous referees for this journal for very valuable comments on earlier drafts. Finally, I thank the Swiss National Science Foundation for generous support of this work (grant 1114-066803.01/1). Please address correspondence to Michael Baumgartner, University of Bern, Department of Philosophy, Laenggassstrasse 49a, 3012 Bern, Switzerland; email: baumgarnter@philo.unibe.ch.

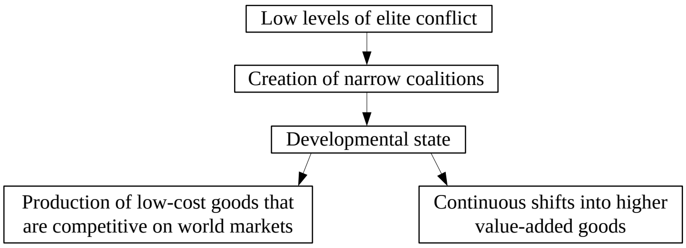

Figure 1. A complex structure consisting of a chainlike and a common cause substructure that cannot be directly analyzed by QCA.

still too large and too complex for an in-depth qualitative analysis. QCA, accordingly, is designed so that it occupies a middle ground between the variable-oriented and the case-oriented traditions. QCA as presented in Ragin (1987) treats single cases in its input data as complex configurations of dichotomous variables. Cases feature one dependent (effect or outcome) variable and an arbitrary amount of independent (possible cause) variables. By a systematic comparison—implementing Boolean techniques—of such configurations, conjunctions of the independent variables can be identified as complex causes of the dependent variable. Every dependent variable can have several alternative complex causes which are disjunctively concatenated in the output of QCA. Complex causes are seen as sufficient conditions, disjunctions of alternative causes as necessary conditions of their effects. The Boolean techniques are, inter alia, resorted to in order to minimize complex conditions that involve redundant variables.

Dichotomous variables correspond to conventional crisp sets. In Ragin (2000) and Ragin (2008), QCA has been thoroughly (and fruitfully) adapted for fuzzy sets, yet the fundamental presumptions of QCA and, most of all, the significant limitations on the complexity of the causal structures that can be uncovered by QCA have remained unaltered: QCA is designed to analyze causal structures featuring exactly *one* effect and a possibly complex configuration of *mutually independent direct* causes of that effect.[1] For brevity, call the assumed singularity of the analyzed effect the *singularity assumption*, or (SNG) for short, and the assumed mutual independence of causes the *independence assumption*, or (IND) for short. Furthermore, an application of QCA always presupposes that it is *known*—or at least determinable prior to implementing QCA—what variable within the analyzed set of variables is the effect, and, accordingly, what variables are possible causes. I shall refer to the assumed *identifiability* of causes and effects as (ICE).[2]

Certain ubiquitous causal structures violate (SNG) and (IND), most notably causal chains and common cause structures. Indeed, it is fair to say that the majority of actual causal structures—also in the area of social sciences (cf. Goertz 2006)—by far exceed the complexity allowed by (SNG) and (IND). For example, consider Waldner's (1999:9) causal models of the connection between state building and economic development in Turkey, Syria, Taiwan, and Korea as graphed in figure 1: Low levels of elite conflict are necessary causes of the creation of narrow coalitions which are necessary for a developmental state. The latter, in turn, is a sufficient cause of two parallel effects: production of low-cost goods that are competitive on world markets, on the one hand, and continuous shifts into higher value-added goods, on the other (cf. also Mahoney and Goertz 2006). This structure, hence, is built up of two causal chains and one common cause substructure. It features both multiple effects and mutually dependent causes, i.e. it violates (SNG) and (IND). Of course, a subdivision of such a complex structure into its separate layers—e.g. 'creation of narrow coalitions' $\longrightarrow$ 'developmental state' or 'developmental state' $\longrightarrow$ 'production of low-cost goods'—yields causal substructures that satisfy both (SNG) and (IND). Such a subdivision would thus render a complex structure amenable to a stepwise QCA analysis: First, the dependency between the creation of narrow coalitions and a developmental state and then the one between a developmental state and the production of low-cost goods could be uncovered by means of QCA. However, such a breaking down of complex structures into simple ones that satisfy (SNG) and (IND) presupposes that a great deal about the very structure under investigation is known prior to its analyzability by QCA—hence (ICE). The variables involved in the investigated cases must be categorized into possible causes and possible effects prior to implementing QCA. In the end, what QCA determines is merely whether possible causal dependencies in fact exist and whether the cause variables constitute complex or alternative causes of the effect or outcome under consideration.

QCA draws on ideas developed within the regularity theoretic tradition of the philosophy of causation. Ragin himself sees QCA as a generalization and systematization of Mill's methods of agreement and difference, and the core of the Boolean techniques to minimize causal conditions used by QCA can be found in Broad (1930), Broad (1944), or Mackie (1974).[3] The paper at hand intends to show that prior knowledge about the causal structure under investigation does not need to be presupposed against a regularity theoretic background. The latter allows for causal reasoning without presuming (SNG) and (IND). Thus, in what follows, I shall present a procedure of causal reasoning that processes the same kind of empirical data as QCA, yet does neither presuppose (SNG) nor (IND) nor (ICE). As will become apparent as we proceed, abandoning these assumptions induces an adaptation of the implemented Boolean techniques. While QCA essentially rests on the well-known Quine-McCluskey optimization of truth functions, the procedure to be presented here significantly deviates from the Quine-McCluskey algorithm.

Clearly, in the course of causally analyzing concrete processes in social sciences prior causal knowledge is often available that guarantees a satisfaction of either of the assumptions (SNG), (IND), or (ICE). In that case, of course, such prior knowledge must be resorted to in order to simplify causal analyses. The more that is known about a causal structure prior to analyzing it, the easier it is successfully uncovered. One of the core motivations behind the procedure advanced in this paper is simply to *minimize* the amount of prior causal knowledge necessary to derive causal structures from small-N data and to show that, in the end, such data can be causally interpreted *without* prior knowledge about the very structure under investigation, irrespective of the latter's complexity. Contrary to QCA, the procedure presented here directly uncovers *causal chains* and *common cause structures*.

2. THE BACKGROUND

Prior to introducing its algorithmic steps, the regularity theoretic background of the inference procedure to be developed in this paper shall be very briefly reviewed.[4] Regularity theories of causation analyze causes and effects on type level, i.e. event types—or *factors* for short—are seen as the primary relata of the causal relation. A factor that causes another factor is said to be *causally relevant* to the latter. Factors are taken to be similarity sets of event tokens, i.e. sets of type identical events or occurrences. Whenever a member of such a similarity set occurs, the corresponding factor is said to be *instantiated*. Factors are symbolized by italicized capital letters A, B, C, etc., with variables Z, Z_1, Z_2, etc. running over the domain of factors. They are negatable. The negation of a factor A is written thus: $\overline{A}$. $\overline{A}$ is simply defined as the complementary set of A. Alternatively, factors can be seen as binary variables that take the value 1 whenever an event of the corresponding type occurs and the value 0 whenever no such event occurs.[5]

Causal analyses are always relativized to a set of investigated factors. This set is referred to as the *factor frame* of the analysis. Factors are virtually never causally relevant to their effects in isolation. Rather, they are parts of whole causing complexes—*complex causes*. A complex cause only becomes causally effective if all of its constituents are co-instantiated, i.e. instantiated coincidently. Coincidently instantiated factors are termed *coincidences*. As will be shown below, coincidences constitute the empirical data processed by the procedure developed in this paper.[6]

Essentially, modern regularity theories analyze causal relevance with recourse to minimalized regularities among factors. The crucial notion needed in the definiens of causal relevance is the notion of a *minimal theory*. Briefly, a minimal theory of a factor B is a *minimally necessary* disjunction of *minimally sufficient* conditions of B. A conjunction of coincidently instantiated factors $A_1 \wedge A_2 \wedge \ldots \wedge A_n$, which for simplicity shall be abbreviated by a mere concatenation of the respective factors, is a minimally sufficient condition of a factor B iff $A_1 A_2 \ldots A_n$ is sufficient for B, i.e. $A_1 A_2 \ldots A_n \rightarrow B$, and there is no proper part α of $A_1 A_2 \ldots A_n$ such that $\alpha \rightarrow B$. A "proper part" of a conjunction designates the result of any reduction of this conjunction by one conjunct. Analogously, a disjunction of factors $A_1 \vee A_2 \vee \ldots \vee A_n$ is a minimally necessary condition of a factor B iff $A_1 \vee A_2 \vee \ldots \vee A_n$ is necessary for B, i.e. $B \rightarrow A_1 \vee A_2 \vee \ldots \vee A_n$, and there is no proper part β of $A_1 \vee A_2 \vee \ldots \vee A_n$ such that $B \rightarrow \beta$. A "proper part" of a disjunction designates the result of any reduction of this disjunction by one disjunct.

That a disjunction of minimally sufficient conditions of a factor B is minimally necessary for B shall be symbolized by "$\Rightarrow$", which is termed a *double-conditional*. Thus, the following is an exemplary minimal theory:

$$AC \vee DE \vee FGH \Rightarrow B \tag{1}$$

Informally, (1) says that whenever AC or DE or FGH are instantiated, B is instantiated as well; that whenever B is instantiated, AC or DE or FGH are instantiated as well; and that sufficient and necessary conditions contained in (1) are minimal. Membership in a minimal theory induces *direct* causal relevance: A factor A is directly causally relevant to a factor B iff A is part of a minimal theory of B. Hence, (1) represents a causal structure such that AC, DE, and FGH are alternative (direct) complex causes of B.

Analyzing the disjunction of alternative causes of B as necessary condition of B amounts to claiming sufficiency of B for just that disjunction. As is often done by critics of regularity accounts, the question might thus be raised as to how the above analysis of causal relevance captures the undisputed non-symmetry of that relation. For if B can be shown to be minimally sufficient for $AC \vee DE \vee FGH$, it might be argued that—relative to the above analysis—B is likewise to be considered causally relevant to its alternative causes. Contrary to first appearances, however, double-conditionals as (1) are not symmetrical with respect to the expressions to the left and the right of "$\Rightarrow$". 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. B only determines the whole disjunction of minimally sufficient conditions. AC and DE and FGH are each minimally sufficient for B, the latter however is only minimally sufficient for $AC \vee DE \vee FGH$. This non-symmetry corresponds to the direction of determination.

Accounting for the non-symmetry of causal relevance in this vein has an important implication as regards the minimal complexity of causal structures. A condition AC that is both minimally sufficient and necessary for a factor B cannot be identified as cause of B, for B is minimally sufficient and necessary for AC as well. All empirical evidence such mutual dependencies generate are perfectly correlated instantiations of AC and B—both are either co-instantiated or absent. Such empirical data can only be causally interpreted if external non-symmetries—as e.g. temporal order—among the instances of AC and B are available. However, as the procedure to be presented in this paper shall infer causal structures on the same empirical basis as QCA, i.e. on the basis of mere coincidence information, perfect correlations among factors shall be taken not to be causally interpretable in the present context. In order to distinguish causes from effects and to orient the cause-effect relation based on coincidence information alone, at least *two* alternative causes are needed for each effect.[7]

Ordinary causal structures far exceed (1) in complexity. Most causally relevant factors are of no interest to causal investigations or are unknown. That is why minimal theories must be either relativized to a specific causal background or kept open for later extensions. The latter is achieved by means of variables. Variables $X_1, X_2, \ldots$ are introduced to stand for an open (but finite) number of additional conjuncts within a sufficient condition, while $Y_A, Y_B, \ldots$ are taken to stand for an open number of additional disjuncts in a minimal theory. If (1) is in this sense kept open for additional factors, one gets:

$$ACX_1 \vee DEX_2 \vee FGHX_3 \vee Y_B \Rightarrow B \tag{2}$$

While direct causal relevance is analyzed with recourse to membership in *simple* minimal theories as (1) and (2), complex causal structures as chains or common cause structures are represented by *complex* minimal theories. Simple minimal theories can be conjunctively concatenated to complex theories: A conjunction of two minimal theories Φ and Ψ is a complex minimal theory iff, first, at least one factor in Φ is part of Ψ and, second, Φ and Ψ do not have an identical consequent.[8] The following are two exemplary complex minimal theories:

$$(AX_1 \vee DX_2 \vee Y_B \Rightarrow B) \wedge (BX_4 \vee GX_5 \vee Y_H \Rightarrow H) \tag{3}$$

$$(AX_1 \vee DX_2 \vee Y_B \Rightarrow B) \wedge (DX_4 \vee GX_5 \vee Y_H \Rightarrow H) \tag{4}$$

(3) represents a causal chain—B is the effect factor of the first conjunct and a cause factor in the second conjunct—, (4) stands for a common cause structure—D is the common cause of B and H. In this vein, causal structures of arbitrary complexity can be represented in regularity theoretic terms. Accordingly, a factor A can be said to be *indirectly* causally relevant to a factor B iff there is a sequence of factors $Z_1, Z_2, \ldots, Z_n$, $n \geq 3$, such that $A = Z_1$, $B = Z_n$, and for each i, $1 \leq i < n$: Z_i is part of the antecedent of a simple minimal theory of Z_{i+1}.

3. THE BASIC IDEA AND INPUT DATA

Minimal theories represent causal structures in a transparent way. Conjunctions in the antecedent of a minimal theory stand for complex causes of the factor in the consequent, disjunctions for alternative causes. Hence, minimal theories are directly causally interpretable. Moreover, minimal theories impose very specific constraints on the behavior of the factors contained in them. For instance, (1) says that whenever AC is instantiated, there also is an instance of B. That is, according to (1) the coincidence $AC\overline{B}$ does not occur. Correspondingly, information about occurring and non-occurring coincidences allows for conclusions as to the minimal theory representing the underlying causal structure. If it is known that AC is never realized in combination with $\overline{B}$, while both $A\overline{C}\overline{B}$ and $\overline{A}C\overline{B}$ are observed, it follows that AC is minimally sufficient for B. In this sense, minimal theories constitute the link between the empirical behavior of the factors in an investigated frame and the causal structure behind that behavior. The empirical behavior of the factors allows for inferring minimal theories that describe that behavior, and these minimal theories, in turn, are causally interpretable.

#	U	L	E
c_1	1	1	1
c_2	1	0	1
c_3	0	1	1
c_4	0	0	0

(a)

#	U	L	E
c_1	1	1	1
c_2	1	1	0
c_3	0	1	1
c_4	1	0	1
c_5	1	0	0
c_6	0	1	0
c_7	0	0	1
c_8	0	0	0

(b)

#	U	L	E
c_1	1	1	1
c_2	0	1	1
c_3	1	0	1
c_4	1	0	0
c_5	0	1	0
c_6	0	0	1
c_7	0	0	0

(c)

#	U	L	E
c_1	1	1	1
c_2	0	0	0

(d)

Table 1. Simple examples of coincidence lists as processed by CNA*.*

The inference procedure to be developed here operates on the same data as QCA: coincidences of the factors involved in a causal process whose structure is to be revealed. Accordingly, the procedure shall be termed *coincidence analysis* or CNA for short.[9] Contrary to QCA, however, the data fed into CNA are not required to mark one factor as the effect or outcome. Based on its input data, CNA simply determines for each factor Z_i in the analyzed frame involving, say, n factors which dependencies hold between Z_i and the other $n-1$ factors in the frame. Most of these dependencies will turn out not to be causally interpretable. The possibly causally interpretable dependencies are then minimalized and expressed in terms of minimal theories, which, finally, are straightforwardly causally interpretable as shown above. Moreover, CNA does not require the $n-1$ other factors to be independent, i.e. to be co-instantiable in all 2^{n-1} logically possible combinations.

To illustrate the input data of CNA and to introduce its algorithmic steps, I shall subsequently draw on hypothetical exemplary studies conducted to investigate the causal dependencies among the following factors: 'strong unions' (abbreviated by U), 'strong left parties' (L), and 'high overall level of education' (E). Table 1 contrasts the data collected in four such hypothetical studies. As in case of QCA, the data processed by CNA is listed analogously to truth tables. Tables as in 1 are referred to as *coincidence lists*. The rows in a coincidence list report the different types of cases observed in a respective study. The cases or rows are numbered by c_1, c_2, etc. In coincidence lists a '1' in the column of, say, factor U represents an instance of U, i.e. a country that has strong unions, a '0' in that same column symbolizes the absence of such an instance, i.e. a country with weak unions.[10] Columns of coincidence lists thus record instances and absences of the factor mentioned in the title row, while the rows following the title row specify coincidences of the factors in the title row. For example, the first row of (a) records a country that has strong unions, strong left parties, and a high overall level of education (ULE);

the following row represents a country that has strong unions, weak left parties, and a high level of education ($U\overline{L}E$).

List (a) in table 1 clearly manifests dependencies among its factors. For instance, there is no row in (a) featuring $UL\overline{E}$. That is, relative to list (a), having strong unions and strong left parties is sufficient for a high overall level of education: $UL \to E$. Likewise, there is no row in (a) featuring L in combination with $\overline{E}$, which amounts to the sufficiency of strong left parties for a high level of education. The sufficient condition UL hence contains a sufficient proper part, L, and, accordingly, is not minimally sufficient. Factor L, on the other hand, does not have any sufficient proper parts and thus is minimally sufficient for E. Analogously it can be shown that having weak left parties and a high level of education ($\overline{L}E$) is minimally sufficient for having strong unions (U) in list (a). As will be shown below, some of these dependencies are causally interpretable; others are not.

In contrast, list (b) contains all 8 logically possible configurations of the 3 factors in its frame. (b) is therefore referred to as a *complete* coincidence list. Complete lists do not feature dependencies among their factors. Accordingly, complete lists do not need to be analyzed for dependencies to begin with. Dependencies emerge only in incomplete lists, i.e. in lists that feature less than 2^n coincidences of the n factors in their frame. Upon investigating processes with hard-to-control causal backgrounds, however, all logically possible factor combinations are no rare empirical result in scientific practice. In such cases, it is often possible to exclude certain configurations as "don't care" cases based on prior causal knowledge (cf. Ragin 1987:113–118). Alternatively, significance levels may be introduced that exclude rarely found configurations from consideration (cf. Ragin 2000:109–115). Thus, several methodologies are available that reduce complete coincidence lists such as to render them causally interpretable after all.

List (c) in table 1 can be seen as a reduction of (b) by one row. There is no row in (c) representing a country that has both strong unions and strong left parties but does not have a high level of education. That is, UL is minimally sufficient for E relative to list (c). Finally, list (d) is incomplete as well. It is incomplete to such an extent that *too many* dependencies emerge. According to list (d), every factor is minimally sufficient and necessary for every other factor in the corresponding frame. Such an abundance of dependencies is not causally interpretable, for causes and effects cannot be distinguished. As the previous section has shown, if causal dependencies are to be oriented on the basis of mere coincidence data—and not, as in case of QCA, by assumption (ICE)—, at least two alternative causes are required for each effect. List (d) is a case of what Ragin calls *limited diversity* (cf. e.g. Ragin 2000:139–141, 198–202). As in case of complete lists, prior causal knowledge may provide a means to causally analyze data featuring limited diversity. Based on such knowledge, lists as (d) may be supplemented by additional rows representing coincidences that, notwithstanding the fact that they have not been observed in a given study, are known to be empirically possible. Such methods of data adjustment, however, are not going to be further discussed here.

4. PRESUPPOSITIONS

While CNA dispenses with assumptions (SNG), (IND), and (ICE), it still rests on two important presuppositions, both of which must be adopted, in one form or another, by any procedure of causal reasoning: First, maximal causal interpretability of empirical data is guaranteed only if that data are exhaustive, and, second, the causal background against which empirical data are collected must be homogeneous. Let us take these presuppositions in turn.

Thoroughly uncovering a causal structure requires an exhaustive collection of empirical data generated by that structure. Accordingly, probabilistic procedures of causal reasoning, for example, presume the availability of probability distributions over all exogenous variables, or QCA relies on the realizability of all 2^n configurations of n investigated cause variables. Nonetheless, assumptions as regards the exhaustiveness of empirical data are hardly ever made explicit in studies on causal reasoning.[11] Such an implicit taking for granted of the suitability of input data, however, will not do for the present context. As the previous section has shown, dependencies among n factors emerge only if not all 2^n coincidences are contained in an analyzed list. Of course, however, coincidences may be missing from coincidence lists due not only to causal dependencies among respective factors. Exhaustive data collection may fail for a host of different reasons. Financial or technical resources may happen to be limited in experimental sciences or nature may be found not to provide sufficient data in non-experimental disciplines. Inexhaustive data are likely to be one of the main reasons for hampered causal interpretability of that data. Proper data collection, however, is not part of causal reasoning, but a precondition thereof. That is why (PEX) is endorsed in the present context, which is concerned with matters of causal reasoning only.

Principle of Empirical Exhaustiveness (PEX): CNA-processed data are exhaustive. That is, all coincidences of the analyzed factors that are compatible with the causal structure regulating the behavior of these factors are contained in a CNA-processed coincidence list.

(PEX) guarantees that whenever a coincidence is missing from a CNA-processed list, this is due to underlying causal dependencies. Clearly, (PEX) constitutes a sweeping idealization with respect to data collection. Such an idealization, however, may prove to be useful in many practical contexts. It can be implemented as a gauge by means of which concrete data collections can be measured and thus evaluated. For clearly, if there is reason to believe that a particular study did not collect all the relevant data about an investigated structure and if there is no other source available that supplements missing data, the pertaining structure simply cannot be fully uncovered. Nonetheless, while (PEX) is a precondition of the *maximal* causal interpretability of coincidence data, even inexhaustive data provide some information as to underlying causal structures. For instance, in list (a) of table 1, factors U and L are independent, i.e. all logically possible configurations of U and L are exhibited in (a). This independence will remain unaltered irrespective of additional coincidences introduced into list (a) in the course of further data collection. Thus, even if (a) violates (PEX), U and L can still be inferred to be causally independent in the structure underlying the behavior of the factors in that list. Since we are subsequently going to be concerned only with maximally causally analyzing coincidence data, (PEX) shall be presumed in the following.[12]

Apart from (PEX), an application of CNA must assume that the causal background of an analyzed coincidence list is *causally homogeneous*, i.e. that the behavior of the factors in the investigated frame is *not confounded* by causally relevant factors not contained in the frame. Each analysis of a causal process is limited to a small subset of all factors actually involved in that process. Causal processes are extremely complex. Ordinarily, only a few factors are of interest in the course of a concrete study. Therefore, a coincidence list over a frame consisting of $Z_1, \ldots, Z_n$ must be assumed to be homogeneous with respect to *confounders* not contained in $\{Z_1, \ldots, Z_n\}$. Roughly, if Z_n is an effect, a confounder of Z_n is a sufficient cause X_j of Z_n such that X_j is located on a causal path leading to Z_n that does not contain any of the factors $Z_1, \ldots, Z_{n-1}$. A confounder is a factor or a conjunction of factors by means of which the investigated effect can be manipulated independently of the factors in the frame.[13]

The notion of a confounder is to be understood relative to a corresponding effect. Basically, any factor in an analyzed frame can be seen as effect of an underlying structure. However, as will be shown below, there are several constraints subject to which a factor can be excluded from the set W of potential effects prior to causally analyzing a factor frame. Still, depending on the specific $Z_i \in \mathsf{W}$ analyzed in the course of a particular run of CNA, different factors are to be seen as confounders and, accordingly, must be homogenized. Generally: Input data processed by CNA are assumed to be generated against causally homogeneous backgrounds in the sense of (HC):[14]

Homogeneity (HC): The background of a causally analyzed list of m coincidences over a factor frame containing the set W of potential effects is causally homogeneous iff for every confounder X_j of every factor in W: X_j is absent in the background of one coincidence iff X_j is absent in the backgrounds of all other $m - 1$ coincidences.

While only homogeneous coincidence lists are causally analyzable, (HC) does not guarantee the causal analyzability of coincidence lists. Rather, (HC) prevents causal fallacies. Therefore, a coincidence list may well be homogeneous in terms of (HC), even though confounders are instantiated in its background—as long as these confounders are instantiated in the backgrounds of all coincidences. If confounders are universally instantiated, effects will be present in all coincidences, irrespective of whether the other factors in the frame are present or absent. In this case, no dependencies emerge; thus, no inferences as to underlying causal structures are drawn. As a consequence, no causal fallacies are committed either.

(HC) excludes a number of coincidence lists from causal analyzability. The lists fed into CNA may well reveal certain backgrounds to be causally inhomoge-

#	U	L	E
c_1	1	1	1
c_2	1	0	1
c_3	0	1	1
c_4	1	1	0

(a)

#	U	L	E
c_1	1	0	0
c_2	0	1	0
c_3	0	0	1
c_4	0	0	0

(b)

Table 2. Two coincidence lists that cannot be causally analyzed, for none of the involved factors can be interpreted as effect of an underlying structure in accordance with (HC).

neous. Consider, for instance, the lists in table 2 which are to be seen as reporting the data collected by two further hypothetical studies investigating the causal dependencies among 'strong unions' (U), 'strong left parties' (L), and 'high level of education' (E). Assume L to be an effect of the structure generating list (a) in table 2. A comparison of the cases c_1 and c_2 recorded in that list shows that, if L in fact were the effect of the underlying structure, (a) would violate (HC). The only factor varying in c_1 and c_2 is L; no other factor in the frame $\{U, L, E\}$ is accountable for that variation of L, therefore, it must be due to a varying confounder of L in the unknown or unconsidered background of list (a). That means interpreting L in terms of an effect contradicts the homogeneity assumption. If L is taken to be a cause factor of the underlying structure, (HC) is not violated. Thus, assuming (HC) to hold for list (a) implies that 'strong left parties' cannot be seen as a possible effect or outcome. The same holds for the other two factors in $\{U, L, E\}$. In c_1 and c_3, U is the only varying factor, while no other factor, apart from E, varies in c_1 and c_4. Hence, there is no factor in list (a) that could possibly be an effect of an underlying causal structure in accordance with (HC). Analogous considerations apply to list (b). In c_1 and c_4 of that list, U is the only varying factor, c_2 and c_4 exclude L from being interpretable as an effect, and c_3 and c_4 refuse E admittance into the set of possible effects due to a violation of (HC).

That means there cannot be a causal structure underlying either list (a) or (b) of table 2 that would be compatible with (HC). In neither list there is a factor that could be seen as an effect in accordance with (HC), i.e. $\mathsf{W} = \emptyset$. Whenever for *every* factor Z_i contained in the factor frame of a coincidence list $\mathcal{C}$ there are two cases c_k and c_l in $\mathcal{C}$ such that Z_i is the only factor varying in c_k and c_l, the background against which the data in $\mathcal{C}$ are collected cannot be homogeneous, for there is no causal structure that could possibly generate $\mathcal{C}$ and accord with (HC). I shall in this context speak of *inhomogeneous coincidence lists*. (HC) excludes all inhomogeneous coincidence lists from being processed by CNA.[15] It must be emphasized, however, that the homogeneity of coincidence lists is an assumption to which every inference of CNA must be relativized. It might well be that a list which is not inhomogeneous in the sense defined above, as e.g. list (a) in table 1, in fact is the result of an uncontrolled variation of background confounders. In this sense, only a sufficient and no necessary condition for the inhomogeneity of a coincidence list is given above. Causal inferences drawn by CNA will always be of the form "Given that (HC) is satisfied, such and such must be the underlying causal structure". Homogeneity is never beyond doubt. Rather, depending on the design of a concrete study, the homogeneity of collected data is more or less plausible. The inferences drawn by CNA—or by any other procedure of causal reasoning—are only as reliable as the unconfoundedness of analyzed data.

5. IDENTIFICATION OF POTENTIAL EFFECTS

After having clarified the presuppositions on which CNA rests, we now proceed to introduce its inference rules. As anticipated in the previous section, a first algorithmic step consists of parsing through the factor frame of a coincidence list in order to determine which of the factors could possibly operate as effects within the causal structure to be uncovered. This step yields a set W of factors whose dependencies on the other factors in the corresponding frame are then successively determined by CNA. The identification of potential effects shall not be considered a proper part of CNA, for any sort of context-dependent empirical information or even prior causal knowledge is allowed to enter the determination of W. For instance, if a factor Z_i is generally instantiated *temporally before* every other factor in an analyzed frame $\{Z_1, \dots, Z_n\}$, Z_i cannot function as an effect within the underlying structure. Or *prior causal knowledge* could be available that establishes the members of a proper subset of $\{Z_1, \dots, Z_n\}$ as *root factors*, i.e. as factors that are causes, but no effects within an investigated structure. In both cases there is no need to integrate respective factors in W. CNA does not have to evaluate dependencies among factors that can be excluded from the set of potential effects to begin with. These pragmatic constraints are not systematizable, or, at least, a systematization shall not be attempted here. Accordingly, no recursively applicable or computable rule can be provided, which essentially is why the determination of W is not seen as a proper part of CNA.

Still, the determination of W is not only regulated by spatiotemporal peculiarities of an analyzed process or by prior causal knowledge. As the previous section has shown, factors can be excluded from the set of potential effects based on homogeneity considerations: For a factor Z_i to be a potential effect, it must not be the case that the corresponding coincidence list reports two cases such that Z_i is the only varying factor in those cases.

These considerations yield the following standard as regards the determination of W. To indicate that the non-computable identification of the set of potential effects is a precondition of launching CNA, yet not a proper part thereof, it shall be referred to as "step 0*".

Step 0* – Identification of potential effects: Given a coincidence list $\mathcal{C}$ over a factor frame $\{Z_1, \ldots, Z_n\}$, identify the subset $\mathsf{W} \subseteq \{Z_1, \ldots, Z_n\}$ such that for every Z_i: $Z_i \in \mathsf{W}$ iff

(1) The totality of available information as to the spatiotemporal ordering of the instances of the factors in $\{Z_1, \ldots, Z_n\}$ and the available prior causal knowledge about the behavior of the factors in $\{Z_1, \ldots, Z_n\}$ does not preclude Z_i to be an effect of the underlying causal structure.
(2) $\mathcal{C}$ does not report two cases c_k and c_l such that Z_i is the only factor varying in c_k and c_l.

6. IDENTIFICATION AND MINIMALIZATION OF SUFFICIENT CONDITIONS

After identifying a non-empty set of potential effects, CNA proper sets in. In a first stage, sufficient conditions for each member of W are identified and minimalized. To illustrate this first stage, let us expand the factor frame we considered in previous sections by two factors such that we now investigate the causal dependencies subsisting among the following five factors: 'strong unions' (U), 'high income disparity' (D), 'strong left parties' (L), 'high GNP' (G), 'high overall level of education' (E). Assume that we have no prior causal knowledge about the structuring of the dependencies among these factors. Thus, prima facie, every factor in the frame $\{U, D, L, G, E\}$ can function as a cause or an effect of the underlying structure. Moreover, no factors in our exemplary frame shall be excluded from effect position by additional information as to the spatiotemporal ordering of their instances. Suppose that in this epistemic situation a study is conducted to the effect that each country included in the study instantiates one of the eight cases listed in table 3. Our hypothetical study, hence, shall be assumed to yield the data listed in table 3.

case #	(strong unions) U	(income disparity) D	(strong left) L	(high GNP) G	(high education) E
c_1	1	1	1	1	1
c_2	1	1	1	0	1
c_3	1	0	1	1	1
c_4	1	0	1	0	1
c_5	0	1	1	1	1
c_6	0	1	1	0	1
c_7	0	0	0	1	1
c_8	0	0	0	0	0

Table 3. Exemplary coincidence list to be analyzed by CNA.

For reasons of compatibility with (HC), 'strong unions', 'high income disparity', and 'high GNP' can be excluded from the set W of potential effects. For each of these factors there is a pair of cases in table 3—$\langle c_1, c_5\rangle$ for U, $\langle c_1, c_3\rangle$ for D, $\langle c_1, c_2\rangle$ for G—such that the respective factor is the only varying factor. In consequence, interpreting one of these factors to be an effect of the underlying structure would contradict CNA's homogeneity assumption. 'Strong left parties' and 'high level of education' thus are the only potential effects of the structure generating table 3, i.e. $\mathsf{W} = \{L, E\}$. For each of the factors in W minimally sufficient conditions are now identified. This is done in four steps: (1) a factor $Z_i \in \mathsf{W}$ is selected, (2) sufficient conditions of Z_i are identified, (3) these sufficient conditions are minimalized, and (4) the procedure is restarted at (1) by selecting another $Z_j \in \mathsf{W}$, until all factors in W have been selected. Let us take these steps in turn.

Step 1 – Selection of a potential effect: Randomly select one factor $Z_i \in \mathsf{W}$ such that Z_i has not been selected in a previous run of steps 1 to 4. Z_i is termed *effect**, the factors in $\{Z_1, \ldots, Z_{i-1}, Z_{i+1}, \ldots, Z_n\}$ are referred to as *residuals*.[16]

Step 2 – Identification of sufficient conditions: Identify all sufficient conditions of the effect* Z_i according to the following rule:

(SUF) A coincidence X_k of residuals is *sufficient* for Z_i iff the input list $\mathcal{C}$ contains at least one row featuring $X_k Z_i$ and no row featuring $X_k \overline{Z_i}$.

The order of selecting effects* in step 1 does not matter, as long as it is guaranteed that, eventually, all members of W are selected. According to (SUF), a coincidence of residuals can be sufficient for an effect* only if it is instantiated at least once. Moreover, a coincidence of residuals contained in the input list is not sufficient for a selected effect* if it is also instantiated in combination with the absence of that effect*.

Let us perform these two steps on our example of table 3 by first selecting 'strong left parties' (L) as effect*. Step 2 identifies six sufficient conditions of L, i.e. there are six coincidences of residuals that conform to (SUF): $UDGE$, $UD\overline{G}E$, $U\overline{D}GE$, $U\overline{DG}E$, $\overline{U}DGE$, $\overline{U}D\overline{G}E$. The case c_1 in the first row of table 3 features the coincidence $UDGE$ in combination with L, and there is no row in table 3 that features $UDGE$ in combination with $\overline{L}$. $UDGE$, thus, is a sufficient condition of L according to (SUF). Analogous considerations apply to the other sufficient conditions mentioned above: c_2 is constituted by $UD\overline{G}E$, c_3 by $U\overline{D}GE$, c_4 by $U\overline{DG}E$, c_5 by $\overline{U}DGE$, and c_6 features $\overline{U}D\overline{G}E$ without either of these conditions being contained in combination with $\overline{L}$ in table 3. Thus, each coincidence of residuals listed in the six cases featuring an instance of 'strong left parties' constitutes a sufficient condition of that factor.

Before sufficient conditions of the remaining effect* E are identified, we proceed to minimalize the sufficient conditions of L.

Step 3 – Minimalization of sufficient conditions: The sufficient conditions of Z_i identified in step 2 are minimalized according to the following rule:

(MSUF) A sufficient condition $Z_1Z_2 \ldots Z_h$ of Z_i is *minimally* sufficient iff neither $Z_2Z_3 \ldots Z_h$ nor $Z_1Z_3 \ldots Z_h$ nor … nor $Z_1Z_2 \ldots Z_{h-1}$ are sufficient for Z_i according to (SUF).

Or operationally put:

(MSUF') Given a sufficient condition $Z_1Z_2 \ldots Z_h$ of Z_i, for every $Z_g \in \{Z_1, Z_2, \ldots, Z_h\}$, $h \geq g \geq 1$, and every h-tuple $\langle Z_{1'}, Z_{2'}, \ldots, Z_{h'} \rangle$ which is a permutation of the h-tuple $\langle Z_1, Z_2, \ldots, Z_h \rangle$: Eliminate Z_g from $Z_1Z_2 \ldots Z_h$ and check whether $Z_1 \ldots Z_{g-1}Z_{g+1} \ldots Z_h\overline{Z_i}$ is contained in a row of $\mathcal{C}$. If that is the case, re-add Z_g to $Z_1 \ldots Z_{g-1}Z_{g+1} \ldots Z_h$ and eliminate Z_{g+1}; if that is not the case, proceed to eliminate Z_{g+1} without re-adding Z_g. The result of performing this redundancy check on every factor contained in $Z_1Z_2 \ldots Z_h$ is a set of minimally sufficient conditions of Z_i.

(MSUF) is nothing but an adaptation of the notion of a minimally sufficient condition as defined in section 2 to the context of coincidence lists. (MSUF'), in turn, can be seen as an operational expression of the analysans of the notion of a minimally sufficient condition implemented in (MSUF). That means (MSUF) might be rephrased as follows: A sufficient condition $Z_1Z_2 \ldots Z_h$ of Z_i is *minimally* sufficient iff it results from an application of (MSUF'). At the price of high computational complexity, the formulation of (MSUF') is kept as simple as possible above. The order in which factors are eliminated from sufficient conditions matters as to the minimalization of such conditions—thus the systematic permutation of elimination orders.[17] In many cases, however, it is not necessary to completely perform that permutation. For instance, if an h-tuple $T_1 = \langle Z_1, \ldots, Z_d, Z_{d+1}, \ldots, Z_h \rangle$ has been minimalized by means of (MSUF') up to element Z_d, that minimalization of T_1 can be taken over for all h-tuples $T_2 = \langle Z_1, \ldots, Z_d, Z_{d+1'}, \ldots, Z_{h'} \rangle$ that coincide with T_1 up to element Z_d without reapplying (MSUF') to T_2. Or suppose it has been found that $X_1 = Z_1 \ldots Z_d$ is a minimally sufficient condition of an investigated effect, and a sufficient condition $X_2 = Z_1Z_2 \ldots Z_h$ containing $Z_1 \ldots Z_d$ is to be minimalized by means of (MSUF'). In that case, it is not effective to minimalize X_2 by first eliminating the factors not contained in X_1, for this elimination order would just yield X_1 again.

Further optimizations of (MSUF') are conceivable but will not be discussed here. More importantly, the intuition behind (MSUF') can be more colloquially captured: Every factor contained in a sufficient condition of Z_i is to be tested for redundancy by eliminating it from that condition and checking whether the remaining condition still is sufficient for Z_i or not. A sufficient condition of Z_i is minimally sufficient iff every elimination of a factor from that condition results in the insufficiency of the remaining condition.

Performing step 3 on our exemplary table is straightforward. Step 2 yielded six sufficient conditions of ‘strong left parties’ (L). For simplicity’s sake, I illustrate the minimalization of these six conditions by means of only two examples. First, take $UDGE$. That this sufficient condition is not minimally sufficient for L is seen by removing, say, G and finding that UDE itself is sufficient for L, for table 3 does not contain a row featuring UDE in combination with $\overline{L}$. UDE still is not minimally sufficient. For instance, both D and E can be removed without sufficiency being lost. Table 3 does not report a case featuring $U\overline{L}$, which induces that ‘strong unions’ (U) is sufficient and, since it is a single factor that does not contain proper parts, *minimally* sufficient for ‘strong left parties’ (L). There are other ways to further minimalize UDE: A removal of U and E still yields a sufficient condition of L. Table 3 does not contain a row featuring $D\overline{L}$. Therefore, ‘high income disparity’ (D) is minimally sufficient for L. Second, let us look at the next sufficient condition of L identified by (SUF). $UD\overline{G}E$ is not minimally sufficient because UD can be removed without sufficiency for L being lost. There is no row in 3 featuring $\overline{G}E$ and $\overline{L}$, which induces that $\overline{G}E$ is sufficient for L. If $\overline{G}E$ is further reduced, sufficiency is lost. c_7 features $E\overline{L}$ and c_8 $\overline{GL}$, which amounts to neither E nor $\overline{G}$ being sufficient for L. ‘Low GNP’ & ‘high level of education’ ($\overline{G}E$), hence, is minimally sufficient for ‘strong left parties’ (L). Minimalizing the other sufficient conditions of L by analogously implementing (MSUF’) does not yield any further minimally sufficient conditions. All in all, therefore, applying step 3 to our exemplary table 3 generates the following three minimally sufficient conditions of ’strong left parties’: ‘Strong unions’ (U), ‘high income disparity’ (D), and ‘low GNP’ & ‘high level of education’ ($\overline{G}E$).

After having identified the minimally sufficient conditions of a first factor $Z_i \in \mathsf{W}$, the same needs to be done for all other effects*. We thus need a loop that brings CNA back to step 1, if not all factors in W have been assigned minimally sufficient conditions yet.

Step 4 – (MSUF)-Loop: If all $Z_i \in \mathsf{W}$ have been selected as effects*, proceed to step 5; otherwise go back to step 1.

Applying this loop to our example yields seven sufficient conditions of ‘high level of education’ (E). Each row featuring E comprises a sufficient condition of residuals: $UDLG$, $UDL\overline{G}$, $U\overline{D}LG$, $U\overline{D}L\overline{G}$, $\overline{U}DLG$, $\overline{U}DL\overline{G}$, $\overline{UDL}G$. For example, c_2 of table 3 is constituted by $UDL\overline{G}$ and there is no row featuring $UDL\overline{G}$ along with $\overline{E}$, or c_3 comprises $U\overline{D}LG$ and no row contains $U\overline{D}LG$ in combination with $\overline{E}$. The sufficiency of the other conditions is analogously demonstrated. Employing (MSUF) or (MSUF’) to minimalize these conditions brings forth four minimally sufficient conditions of ‘high level of education’: ‘strong unions’ (U), ‘high income disparity’ (D), ‘strong left parties’ (L), and ‘high GNP’ (G). The list in table 3 contains no rows featuring either $U\overline{E}$, $D\overline{E}$, $L\overline{E}$, or $G\overline{E}$.

As an overall result of performing the first stage (steps 1 to 4) of CNA on our exemplary study, we have thus identified the following minimally sufficient conditions of the factors in W:

$$U, D, \overline{G}E \quad \text{for} \;\; L,$$
$$U, D, L, G \quad \text{for} \;\; E.$$

Before we move on, emphasis must be put on a major difference between QCA and CNA that becomes apparent at this point. By presupposing that potential causes of an investigated structure are independent (IND), and thus are co-instantiable in all logically possible combinations, QCA can draw on the well-known Quine-McCluskey optimization of truth functions in order to minimalize sufficient conditions (cf. Quine 1952, 1959). As soon as (IND) is dropped, however, Quine-McCluskey optimization no longer eliminates all redundancies. Our exemplary coincidence list in table 3 features a dependency among $U \vee D$ and L. There is no row in table 3 reporting a coincidence of, say, U and $\overline{L}$. Yet, Quine-McCluskey optimization only eliminates redundant conjuncts of sufficient conditions if a respective truth table contains two rows which differ only with respect to presence and absence of that conjunct. Thus, minimalizing the sufficient conditions of E in table 3 along the lines of Quine-McCluskey would not identify, say, U as minimally sufficient condition of E, notwithstanding the fact that table 3 does not contain a coincidence of U and $\overline{E}$. Rendering coincidence lists generated by complex causal structures amenable to a Boolean analysis, accordingly, calls for a custom-built minimalization procedure that differs from a standard Quine-McCluskey optimization insofar as it systematically tests conjuncts Z_g of a sufficient condition X_i for eliminability, irrespective of whether the corresponding coincidence list contains another sufficient condition X_j that only differs from X_i with respect to presence and absence of Z_g.

7. IDENTIFICATION AND MINIMALIZATION OF NECESSARY CONDITIONS

As the famous Manchester Hooters counterexample against Mackie's (1974) INUS-theory of causation[18] demonstrates and as articulated in the analysis of causal relevance given in section 2, minimally sufficient conditions are not generally causally interpretable. Only minimally sufficient conditions that are moreover non-redundant parts of minimally necessary conditions are amenable to a causal interpretation. After having identified minimally sufficient conditions, we thus now proceed to first form necessary conditions of the effects* from their minimally sufficient conditions and then remove redundancies from these necessary conditions. Since factor frames processed by CNA are incomplete with respect to underlying causal structures, i.e. there supposedly will always be many causally relevant factors not contained in input lists, effects* can only be assigned necessary conditions relative to the homogeneous backgrounds of corresponding coincidence lists. This is easily accomplished by disjunctively combining the minimally sufficient conditions of each effect*. In this way, we get one necessary condition relative to an input list $\mathcal{C}$ and its background for each factor $Z_i \in \mathsf{W}$.

Step 5 – Identification of necessary conditions: Identify a necessary condition of each effect* Z_i by disjunctively concatenating Z_i's minimally sufficient conditions according to the following rule:

(NEC) A disjunction $X_1 \vee X_2 \vee \ldots \vee X_h$ of minimally sufficient conditions of Z_i is *necessary* for Z_i iff $\mathcal{C}$ contains no row featuring Z_i in combination with $\neg(X_1 \vee X_2 \vee \ldots \vee X_h)$, i.e. no row comprising $\overline{X_1 X_2 \ldots X_h} Z_i$.

Performed on our example, step 5 issues $U \vee D \vee \overline{G}E$ and $U \vee D \vee L \vee G$ as necessary conditions of L and E, respectively. That means there is no row in table 3 featuring L in combination with neither U nor D nor $\overline{G}E$. Every country included in our hypothetical study that has strong left parties also exhibits one of L's minimally sufficient conditions. Similarly for E: No row of table 3 records a coincidence of E with neither an instance of U nor D nor L nor G. Whenever a country in our study has a high level of education, it also features one of E's minimally sufficient conditions.

Such as to determine whether the minimally sufficient conditions assigned to the effects* at the end of the previous section in fact are non-redundant parts of necessary conditions, these necessary conditions have to be minimalized.

Step 6 – Minimalization of necessary conditions: The necessary conditions of every $Z_i \in \mathsf{W}$ identified in step 5 are minimalized according to the following rule:

(MNEC) A necessary condition $X_1 \vee X_2 \vee \ldots \vee X_h$ of Z_i is *minimally* necessary iff neither $X_2 \vee X_3 \vee \ldots X_h$ nor $X_1 \vee X_3 \vee \ldots X_h$ nor ... nor $X_1 \vee X_2 \vee \ldots \vee X_{h-1}$ is necessary for Z_i according to (NEC).

Or operationally put:

(MNEC') Given a necessary condition $X_1 \vee X_2 \vee \ldots \vee X_h$ of Z_i, for every $X_g \in \{X_1, X_2, \ldots, X_h\}$, $h \geq g \geq 1$, and every h-tuple $\langle X_{1'}, X_{2'}, \ldots, X_{h'} \rangle$ which is a permutation of the h-tuple $\langle X_1, X_2, \ldots, X_h \rangle$: Eliminate X_g from $X_1 \vee X_2 \vee \ldots \vee X_h$ and check whether there is a row in $\mathcal{C}$ featuring Z_i in combination with $\neg(X_1 \vee \ldots \vee X_{g-1} \vee X_{g+1} \vee \ldots \vee X_h)$, i.e. a row comprising $\overline{X_1 \ldots X_{g-1} X_{g+1} \ldots X_h} Z_i$. If that is the case, re-add X_g to $X_1 \vee \ldots \vee X_{g-1} \vee X_{g+1} \vee \ldots \vee X_h$ and eliminate X_{g+1}; if that is not the case, proceed to eliminate X_{g+1} without re-adding X_g. The result of performing this redundancy check on every minimally sufficient condition contained in $X_1 \vee X_2 \vee \ldots \vee X_h$ is a set of minimally necessary conditions of Z_i.

In analogy to (MSUF), (MNEC) is nothing but an adaptation of the notion of a minimally necessary condition as defined in section 2 to the context of coincidence lists. (MNEC'), in turn, can be seen as an operational expression of the analysans of the notion of a minimally necessary condition implemented in

(MNEC). That means (MNEC) might be rephrased as follows: A necessary condition $X_1 \vee X_2 \vee \ldots \vee X_h$ is *minimally* necessary iff it results from an application of (MNEC'). The formulation of (MNEC') has been kept as simple as possible at the expense of its computational complexity. Analogous optimizations as in case of (MSUF') (cf. p. 15 above) are possible with respect to (MNEC'). The intuition behind (MNEC') can also be more colloquially captured: Every minimally sufficient condition contained in a necessary condition of Z_i is to be tested for redundancy by eliminating it from that condition and checking whether the remaining condition still is necessary for Z_i. A necessary condition of Z_i is minimally necessary iff every elimination of a minimally sufficient condition from that necessary condition results in the loss of necessity of the remaining condition.

Let us illustrate step 6 by first performing it on the necessary condition $U \vee D \vee \overline{G}E$ of L. That disjunction is not minimally necessary for L because it contains a necessary proper part: $U \vee D$. Every country with strong left parties included in the study behind table 3 has either strong unions or a high income disparity. Table 3 does not contain a row featuring $\overline{UD}L$. $\overline{G}E$ does not amount to a non-redundant part of a minimally necessary condition, for whenever $\overline{G}E$ is instantiated in combination with L, there also is an instance of $U \vee D$. The same results from applying (MNEC') to $U \vee D \vee \overline{G}E$. When eliminating U, we find that the rest is no longer necessary for L, because c_3 of table 3 features $\overline{D\overline{G}}E$ and L, or more specifically $\overline{D}GE$ and L. Hence, U is re-added. The same is found after removing D. c_5 features $\overline{U\overline{G}}E$ and L or $\overline{U}GE$ and L, respectively. Removing $\overline{G}E$, however, does not result in a loss of necessity. Therefore, $\overline{G}E$ is not re-added. $U \vee D \vee L \vee G$ does not amount to a minimally necessary condition of E either. $U \vee D \vee L \vee G$ contains not only one but two necessary proper parts: $L \vee G$ and $U \vee D \vee G$. There is no row in table 3 featuring $\overline{LG}E$ or $\overline{UDG}E$. Every country in our exemplary study that has a high level of education also features an instance of $L \vee G$ and one of $U \vee D \vee G$. These two ways to minimalize $U \vee D \vee L \vee G$ stem from the fact that there are dependencies among the minimally sufficient conditions of E. Within the homogeneous background of table 3, L is instantiated if and only if $U \vee D$ is instantiated. All in all, therefore, we get the following minimally necessary conditions for our example:

$$\begin{aligned} U \vee D \quad & \text{for} \;\; L, \\ L \vee G \text{ and } U \vee D \vee G \quad & \text{for} \;\; E. \end{aligned}$$

8. FRAMING MINIMAL THEORIES AND CAUSAL INTERPRETATION

Step 6 of CNA yields a set of minimally necessary disjunctions of minimally sufficient conditions for each $Z_i \in \mathsf{W}$. We have thus come close to assigning minimal theories to the data of our hypothetical study. The result of step 6 allows for framing one simple minimal theory for 'strong left parties' and two for 'high level of education'. Relative to the background of table 3, these minimal theories can be

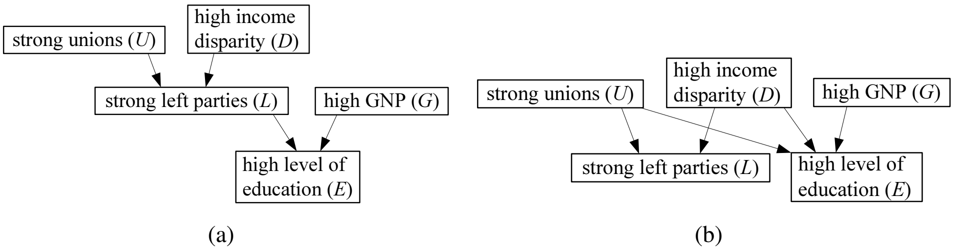

Figure 2. Two complex structures one of which underlies the coincidences in table 3.

straightforwardly expressed as follows: $U \vee D \Rightarrow L$, $U \vee D \vee G \Rightarrow E$, $L \vee G \Rightarrow E$. However, apart from the specific causal background of the particular study represented in table 3, it must not be the case that U and D are themselves sufficient for L or L and G for E. Moreover, there may well be further minimally sufficient conditions of both 'strong left parties' and 'high level of education'. Therefore, suspending the relativization to the background of table 3 and expressing these dependencies in their general and background-independent form leads to:

$$UX_1 \vee DX_2 \vee Y_L \Rightarrow L \tag{5}$$

$$UX_1 \vee DX_2 \vee GX_3 \vee Y_E \Rightarrow E \tag{6}$$

$$LX_1 \vee GX_2 \vee Y_E \Rightarrow E \tag{7}$$

L and E have a non-empty intersection of minimally sufficient conditions. Correspondingly, the simple minimal theories of L and E have certain factors in common. The causal structure regulating E is not independent of the structure behind L. The behavior of the factors in table 3, thus, is regulated by a complex structure. To determine what that structure looks like, the simple minimal theories of L and E are to be conjunctively combined to form a complex theory. Here an ambiguity emerges: (6) and (7)—if causally interpreted—identify different direct causal relevancies for E. While according to (6) U and D are directly causally relevant to E, (7) instead holds L to be directly relevant to E. The coincidences in table 3 are either generated by a causal chain such that U and D are parts of alternative causes of L while L and G are contained in alternative causes of E, or they are generated by a common cause structure such that U and D are parts of alternative causes of L while U, D, and G are contained in alternative causes of E. The two causal structures possibly underlying table 3 are graphed in figure 2. Thus, the minimalization of E's necessary condition is ambiguous.[19]

The data listed in table 3 alone do not determine whether the interplay of 'strong unions', 'high income disparity', 'strong left parties', 'high GNP', and 'high level of education' is regulated by a chain or a common cause structure. If no prior causal knowledge about the structure under investigation is at hand that disambiguates the inference, a disambiguation has to await further study, i.e. expansions of the factor frame and a corresponding collection of additional data. For instance, if a follow-up study reveals a country that has strong left parties and a low level of education, the structure behind list 3 can unambiguously be identified as a common cause structure. As has been widely recognized and explored in the literature concerned with probabilistic causal reasoning, ambiguities with respect to the causal interpretation of empirical data constitute a very common phenomenon in causal reasoning.[20] I systematically investigate the ambiguities that may arise in the course of Boolean causal reasoning in Baumgartner (2008a).

To capture ambiguities of this kind, CNA cannot assign one single minimal theory to our exemplary data. Rather, in order for a Boolean methodology to be adequate it must assign a *set* of minimal theories to its input data, all of which could represent the causal structure that in fact accounts for the data. Therefore, the remaining step of CNA frames all minimal theories that can be constructed from the inventory of minimally necessary disjunctions of minimally sufficient conditions identified for each $Z_i \in \mathsf{W}$ in step 6. This is done by means of a twofold procedure: First, simple minimal theories are formed for each $Z_i \in \mathsf{W}$, and second, if the minimal theories Φ and Ψ of two different factors in W have a non-empty intersection of factors, Φ and Ψ are combined to form the complex minimal theory $\Phi \wedge \Psi$, such that $\Phi \wedge \Psi$ conforms to the requirements imposed on the notion of a complex minimal theory in section 2.

Step 7 – Framing minimal theories: The minimally necessary disjunctions of minimally sufficient conditions of each $Z_i \in \mathsf{W}$ identified in step 6 are assembled to minimal theories as follows:

(1) For each $Z_i \in \mathsf{W}$ and each minimally necessary disjunction $X_1 \vee X_2 \vee \ldots \vee X_h$, $h \geq 2$,[21] of minimally sufficient conditions of Z_i: Form a simple minimal theory Ψ of Z_i by making $X_1 \vee X_2 \vee \ldots \vee X_h$ the antecedent of a double-conditional and Z_i its consequent: $X_1 \vee X_2 \vee \ldots \vee X_h \Rightarrow Z_i$.

(2) Conjunctively combine two simple minimal theories Φ and Ψ to the complex minimal theory $\Phi \wedge \Psi$ iff Φ and Ψ conform to the following conditions:

(a) at least one factor in Φ is part of Ψ;

(b) Φ and Ψ do not have an identical consequent.

Applied to our exemplary study, step 7 frames the minimal theories (8) and (9), which are given both in background dependent and independent form below. (8) represents the chain (a) of figure 2, while (9) stands for the common cause structure (b).

$$\begin{array}{c} (U \vee D \Rightarrow L) \wedge (L \vee G \Rightarrow E) \\ \scriptstyle (UX_1 \vee DX_2 \vee Y_L \Rightarrow L) \wedge (LX_3 \vee GX_4 \vee Y_E \Rightarrow E) \end{array} \tag{8}$$

$$\begin{array}{c} (U \vee D \Rightarrow L) \wedge (U \vee D \vee G \Rightarrow E) \\ \scriptstyle (UX_1 \vee DX_2 \vee Y_L \Rightarrow L) \wedge (UX_1X_3 \vee DX_2X_3 \vee Y_LX_3 \vee GX_4 \vee Y_E \Rightarrow E) \end{array} \tag{9}$$

After having assigned a set of minimal theories to a coincidence list, the by far most intricate hurdles on the way to a causal analysis of table 3 have been overcome. As we have seen in section 2, there exists a straightforward syntactical convention as regards the causal interpretation of minimal theories. Minimal theories render causal structures syntactically transparent:

Step 8* – Causal interpretation: Disjuncts in the antecedents of simple minimal theories are to be interpreted as alternative (complex) causes of the factor in the consequent. Conjuncts constituting such disjuncts correspond to non-redundant parts of complex causes. Triples of factors $\langle Z_h, Z_i, Z_j \rangle$, such that Z_h appears in the antecedent of a minimal theory of Z_i and Z_i is part of a minimal theory of Z_j, are to be interpreted as causal chains.

This interpretation rule is not to be seen as part of CNA proper. Nonetheless, it fulfills an essential function on the way to a causal inference. For this reason, the rule concerning causal interpretation is starred.

All in all, CNA thus determines the coincidences in our exemplary data table 3 to be the result of the causal chain represented by graph (a) of figure 2 or of the common cause structure (b). Steps 0* to 7 assign a set of minimal theories to a coincidence list and step 8* causally interprets these theories.

9. SUMMARY

This paper introduced a procedure of causal reasoning that is embedded in a regularity theoretic framework and implements mainly Boolean techniques. Coincidence analysis (CNA) differs from QCA essentially in three respects: First, CNA does not assume that there is exactly one effect in every causally analyzed factor frame, second, CNA does not presuppose the mutual independence of the causes of that effect, and third, as to CNA it must not be known prior to applying CNA what factor within the analyzed frame is the effect and, accordingly, what factors are possible causes. Thus, CNA abandons the three QCA-assumptions (SNG), (IND), and (ICE). It has been shown that these causal assumptions made in the context of QCA are dispensable for a successful causal analysis of coincidence information. Thus, homogeneity (HC) turns out to be the only causal assumption needed for causal reasoning based on pure coincidence data; and contrary to (SNG), (IND), and (ICE), (HC) is not a causal assumption about the very structure under investigation, but about the latter's causal background.

As an immediate consequence thereof, CNA is not limited to uncovering causal structures layer by layer. While QCA is only applicable provided that prior causal knowledge separates analyzed factor frames in a subset consisting of causally independent (possible) cause factors and a subset consisting of a single effect, CNA is applicable even without any prior causal knowledge concerning the underlying structure. CNA is capable of analyzing causal structures from scratch and in their whole complexity. Due to limited space, of course, the one example discussed here is simple and designed in such a way that the performance of CNA with respect to complex causal structures that are critical for QCA becomes transparent. In principle, however, CNA is capable of analyzing structures of arbitrary complexity. Considerably more complex examples can be found in Baumgartner (2006).

Apart from generalizing QCA, CNA fills a gap left open by the probabilistic algorithms of causal reasoning as presented in Spirtes, Glymour, and Scheines (2000). These algorithms only generate informative outputs provided that analyzed conditional probabilities are lower than 1, i.e. provided that causes do not in a strict sense determine their effects. CNA, in contrast, is custom-built for deterministic causal dependencies and properly uncovers such dependencies.

As shown in sections 3 and 4, not every coincidence list is causally analyzable. Accordingly, CNA cannot be seen as a *complete* inference procedure in the sense that it assigns a causal structure to a coincidence list whenever the coincidences in that list are in fact the result of such a structure. Empirical data may be insufficient to uncover causal regularities. However, CNA is a *correct* causal inference procedure in the sense that whenever CNA assigns a set of causal structures to a coincidence list, that list is in fact generated by a member of that set. CNA assigns sets of minimal theories and thereby sets of causal structures to every causally interpretable coincidence list.

Notes

[1]In fact, Ragin explicitly tailors his notion of *causal complexity* to these limitations on QCA-processable complexity. He defines *causal complexity* as "a situation in which a given outcome may follow from several different combinations of causal conditions" (2008:124), similarly in Ragin (1987:23–26). Plainly, as we shall see below, this is a rather sweeping simplification of the actual complexity of real-life causal structures.

[2]It shall not be claimed that these 3 assumptions are logically independent. They are just labeled here for the purpose of easy reference later on. Moreover, it must be pointed out that (SNG), (IND) and (ICE) are not *explicitly* assumed in the context of QCA, rather they are *implicitly* taken for granted.

[3]Cf. also Quine (1952) and Quine (1959).

[4]For details on the theoretical background resorted to here see Graßhoff and May (2001), Baumgartner and Graßhoff (2004), and Baumgartner (2008b).

[5]That means the procedure developed in this paper is custom-built for causal structures featuring binary variables. This restriction primarily serves conceptual simplicity, as it allows for a straightforward implementation of Boolean optimization techniques. In consequence, structures involving multi-valued variables must be encoded in binary terms before they can be treated by the procedure introduced here. For quite some time, however, there have been considerable efforts in the literature on logic synthesis to generalize Boolean optimization procedures for systems involving multi-valued variables (cf. e.g. Mirsalehi and Gaylord 1986, or Sasao 1999, ch. 10). Hence, there seem to be no principled obstacles to generalizing the procedure introduced here for multi-valued variables as well—possibly by suitably adapting the ideas of Cronqvist and Berg-Schlosser (2008).

[6]Coincidences correspond to what Ragin (1987) calls *configurations*.

[7]QCA does not face the problem of the orientation of causal dependencies, for applying QCA is taken to be possible only after the effect has been identified within the analyzed factor frame (cf. ICE). However, as (ICE) shall be given up here, a way to orient causal dependencies is needed.

[8]The first constraint guarantees that complex minimal theories represent cohering causal structures, and the second restriction prohibits the conjunctive concatenation of equivalent minimal theories and thus excludes redundancies.

[9]Coincidence analysis is not abbreviated as "CA" because, in the social science literature, this acronym is often used for *correspondence analysis* which must not be confused with coincidence analysis.

[10]I shall in the present context sidestep the problems that inevitably arise when it comes to categorizing countries into e.g. the ones with strong and the ones with weak unions. The problems of attributing vague or fuzzy properties are extensively addressed in Ragin (2000, 2008).

[11]One exception is Ragin (1987, 2000, 2008). He discusses at length how insufficiently diverse data negatively affect causal reasoning.

[12]I address the problems arising from violations of (PEX) in Baumgartner (2008c).

[13]For more details on the notion of a confounder needed for the purposes of CNA cf. Baumgartner (2008c).

[14]Ragin is rather vague about the details of the homogeneity assumption adopted by QCA. Ragin et. al. (1996), for instance, stipulate that causal data analysis presupposes that "the cases at the start of an investigation are in fact alike enough to permit comparisons" (752). (HC) can be seen as an attempt to make more precise what "alike enough to permit comparisons" means in the context of Boolean causal reasoning.

[15]In a way, QCA is more permissive with respect to the causal analyzability of inhomogeneous lists. If prior causal knowledge is available that excludes (at least) one case c_i recorded in an inhomogeneous list from consideration—say, because it is known that a causal structure is operating in c_i that differs from the structure under investigation such that c_i is explained away—, QCA could be applied to inhomogeneous lists. Since one of the main motivations behind CNA is to minimize the amount of prior causal knowledge needed to causally interpret small-*N* data, invoking such data adjustments in order to analyze inhomogeneous lists by CNA is uncalled for.

[16]Selected factors are labeled *effects** to indicate that they *possibly* are the effects of the structure generating the input list. Effects* do not necessarily turn out to be (actual) effects of the underlying structure at the end of a CNA-analysis. For instance, the set of effects* contained in list (d) of table 1 contains all factors in the frame—provided no further information is available that distinguishes among causes and effects. Yet, none of these effects* is identified as actual effect by CNA.

[17]This is an important deviation from the minimalization of sufficient conditions as performed by QCA. In the vein of the Quine-McCluskey optimization of truth functions QCA only eliminates conjuncts of a sufficient condition if the latter reduced by the respective conjunct is actually contained in the coincidence list. As will be shown below, this restriction is a serious limitation of the minimizability of sufficient conditions involved in complex causal structures.

[18]Cf. Mackie (1974), Baumgartner and Graßhoff (2004), ch. 5

[19]That the minimalization of necessary conditions can be ambiguous is not taken into account in the context of QCA. In Ragin (1987, 2000, 2008) the minimalization of necessary conditions is assumed to be unproblematic. For another ambiguity with respect to minimalizing necessary conditions cf. Quine (1959) and Kim (1993).

[20]Cf. e.g. Frydenberg (1990), Verma and Pearl (1991), or Spirtes et al. (2000).

[21]The constraint as to a minimum of two alternative minimally sufficient conditions for each effect* does justice to the minimal complexity of a causal structure required such that its direction is identifiable (cf. section 2).

REFERENCES

Baumgartner, M. 2006. *Complex Causal Structures. Extensions of a Regularity Theory of Causation*, Ph.D. dissertation, University of Bern.

———. 2008a. "The Causal Chain Problem." *Erkenntnis* 69:201–226.

———. 2008b. "Regularity Theories Reassessed." *Philosophia* 36:327–354.

———. 2008c. "Uncovering Deterministic Causal Structures: A Boolean Approach." *Synthese* URL: ⟨ http://www.springerlink.com/content/x0487831qk67h455/ ⟩.

Baumgartner, M. and G. Graßhoff. 2004. *Kausalität und kausales Schliessen. Eine Einführung mit interaktiven Übungen*. Bern: Bern Studies Verlag.

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

———. 1944. "Hr. von Wright on the Logic of Induction I-III." *Mind* 53:1–24, 97–119, 193–214.

Cronqvist, L. and D. Berg-Schlosser. 2008. "Multi-Value QCA (mvQCA)." Pp. 69–86 in *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*, edited by B. Rihoux and C. C. Ragin. Thousand Oaks, CA: Sage.

Frydenberg, M. 1990. "The Chain Graph Markov Property." *Scandinavian Journal of Statistics* 17:333–353.

Goertz, G. 2006. *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.

Graßhoff, G. and M. May. 2001. "Causal Regularities." Pp. 85–114 in *Current Issues in Causation*, edited by W. Spohn, M. Ledwig, and M. Esfeld. Paderborn: Mentis.

Kim, J. 1993. "Causes and Events: Mackie on Causation." Pp. 60–74 in *Causation*, edited by E. Sosa and M. Tooley. Oxford: Oxford University Press.

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

Mahoney, J. and G. Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14:227–249.

Mirsalehi, M. M. and T. K. Gaylord. 1986. "Logical Minimization of Multilevel Coded Functions." *Applied Optics* 25:3078–3088.

Quine, W. v. O. 1952. "The Problem of Simplifying Truth Functions." *The American Mathematical Monthly* 59:521–531.

———. 1959. "On Cores and Prime Implicants of Truth Functions." *The American Mathematical Monthly* 66:755–760.

Ragin, C. C. 1987. *The Comparative Method*. Berkeley: University of California Press.

———. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

———. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Ragin, C. C., D. Berg-Schlosser, and G. De Meur. 1996. "Political Methodology: Qualitative Methods." Pp. 749–768 in *New Handbook of Political Science*, edited by R. Goodin and H. D. Klingemann. New York: Oxford University Press.

Sasao, T. 1999. *Switching Theory for Logic Synthesis*. Boston: Kluwer.

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

Verma, T. S. and J. Pearl. 1991. "Equivalence and Synthesis of Causal Models." Pp. 255–268 in *Uncertainty in Artificial Intelligence 6*, edited by P. P. Bonissone, M. Henrion, L. N. Kanal, and J. F. Lemmer. Amsterdam: Elsevier.

Waldner, D. 1999. *State Building and Late Development*. Ithaca: Cornell University Press.

Michael Baumgartner is a lecturer and research associate at the philosophy department of the University of Bern, Switzerland. His publications include an introduction to the philosophy of causation entitled "Kausalität und kausales Schliessen" (2004) and a number of papers on causation, regularity theories, interventionism, causal reasoning, and logical formalization.