Detecting Causal Chains in Small-N Data

Michael Baumgartner

The first part of this paper shows that *Qualitative Comparative Analysis* (QCA)—also in its most recent form as presented in Ragin (2008)—, does not correctly analyze data generated by causal chains. The incorrect modeling of data originating from chains essentially stems from QCA's reliance on Quine-McCluskey optimization to eliminate redundancies from sufficient and necessary conditions. Baumgartner (2009a,b) has introduced a Boolean methodology, termed *Coincidence Analysis* (CNA), that is related to QCA, yet, contrary to the latter, does not eliminate redundancies by means of Quine-McCluskey optimization. The second part of the paper applies CNA to chain-generated data. It will turn out that CNA successfully detects causal chains in small-N data.

KEY WORDS: causal modeling; small-N data; causal chains; Qualitative Comparative Analysis; Coincidence Analysis.

1. INTRODUCTION

Since its first detailed presentation in Ragin (1987), *Qualitative Comparative Analysis* (QCA) has become a widely used methodology to causally model small- and intermediate-N data in the social sciences. While QCA has originally been developed for conventional crisp sets, Ragin (2000, 2008) has fruitfully adapted the method for (purposefully calibrated) fuzzy sets. These recent adaptations have considerably widened QCA's domain of applicability and enhanced the level of precision that can be achieved by QCA analyses. At the same time, they have not altered QCA's computational core. By systematic comparisons of the cases constituting QCA's input data, Boolean combinations of conditions are identified as being sufficient and/or necessary for an outcome. As complex sufficient and necessary conditions typically involve redundancies, they must be rigorously minimized before they are amenable to a causal interpretation (cf. Baumgartner 2008). Both in crisp set QCA ($csQCA$) and in fuzzy set QCA ($fsQCA$) such redundancies are eliminated by means of Quine-McCluskey optimization of truth functions.

In section 2, I show that minimizing causal conditions on the basis of Quine-McCluskey optimization imposes significant constraints on the complexity of the causal structures that can be uncovered by use of QCA; in particular, it prevents QCA from correctly modeling data generated by causal chains. (To avoid complications that are dispensable for the purposes of this paper, I am going to focus on crisp set analyses only.) Eliminating redundancies from causal conditions

I thank Ruedi Epple, Gary Goertz, Charles Ragin, and two anonymous referees for this journal for many helpful comments on earlier versions of this paper. Moreover, I am indebted to the Deutsche Forschungsgemeinschaft (DFG) for generous support of this work (project CAUSAPROBA).

by means of Quine-McCluskey optimization requires that all 2^n logically possible configurations of n conditions are compatible with the causal structure under investigation. However, this requirement is not satisfied if, as in case of causal chains, there are causal dependencies among the n analyzed conditions themselves.

Baumgartner (2009a,b) has introduced a Boolean methodology for the causal analysis of configurational data termed *Coincidence Analysis* (CNA) that is related to QCA, yet does not minimize conditions with recourse to Quine-McCluskey optimization. As a direct consequence, CNA does not need to assume that all 2^n logically possible configurations of n investigated conditions are compatible with underlying causal structures. This, in turn, renders CNA applicable to data stemming from causal chains. Section 3 reviews CNA's alternative minimization procedure. While in Baumgartner (2009a,b) CNA has been introduced as a general Boolean procedure that processes any kind of configurational data, the presentation in section 3 is tailored to social scientific practice. Moreover, the aim of section 3 is to clarify the relevant differences between QCA and CNA, rather than to discuss the computational details of CNA. Finally, section 4 shows that its custom-built minimization procedure enables CNA to successfully model data that result from causal chains.

To render the computational differences between QCA and CNA as transparent and accessible as possible, my discussion will turn on artificially simple hypothetical examples and I am going to sidestep all practical complications that inevitably arise when it comes to applying Boolean methodologies to real-life data. Comparing QCA and CNA with respect to their respective handling of real-life data has to await another occasion (for an application of CNA to more complex data tables cf. Baumgartner 2009a).

2. QCA AND CAUSAL CHAINS

To illustrate the basic analytical techniques of QCA, consider the (truth) table 1 which represents hypothetical country-level data on two causal conditions, strong unions (U) and strong left parties (L), and one outcome, generous welfare state (G) (cf. Ragin 2008, chs. 8, 9). The rows of table 1 can be read as standing for *types* of countries. For instance, c_1 represents countries featuring strong unions,

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

Table 1. Exemplary data table as processed by QCA with U representing "strong unions", L "strong left parties", and G "generous welfare state".

strong left parties, and a generous welfare state, whereas c_2 exhibits countries with strong unions, weak left parties, and a generous welfare state, and analogously for the other rows.

In a first step, QCA identifies sufficient conditions of the investigated outcome—G in our example. In table 1, the conjunction of U and L, which I symbolize by a mere concatenation of U and L, is sufficient for G, i.e. it holds that $UL \rightarrow G$.[1] Or in words: whenever a country has strong unions and strong left parties, it also has a generous welfare state. Rows c_2 and c_3 also feature sufficient conditions of G. All countries with strong unions and weak left parties or weak unions and strong left parties exhibit welfare generosity, i.e. $U\overline{L} \rightarrow G$ and $\overline{U}L \rightarrow G$, where $\overline{L}$ and $\overline{U}$ represent the negations of L and U.

In a second step, QCA eliminates all redundancies from these sufficient conditions. As anticipated in the introduction, this is accomplished on the basis of Quine-McCluskey optimization. To determine whether, say, the (complex) condition UL, which is exemplified in c_1 of table 1, is not only sufficient but also *minimally* sufficient for G, Quine-McCluskey optimization requires parsing the input table to find other rows that accord with c_1 in regard to the outcome and all (atomic) conditions except for one. In table 1, two such rows exist for c_1: c_2 and c_3. In c_2 all conditions are the same as in c_1, except for L which is absent in c_2 and present in c_1. In c_3 all conditions are the same as in c_1, except for U which is absent in c_3 and present in c_1. Both rows c_2 and c_3 also feature sufficient conditions of G. The pair of rows $\langle c_1, c_2\rangle$ shows that U alone (independently of L) is sufficient for G, and the pair $\langle c_1, c_3\rangle$ reveals that L alone is sufficient for G. That is, our first sufficient condition UL contains two sufficient proper parts, *viz.* U and L, where a *proper part of a conjunction* $Z_1Z_2 \ldots Z_n$ designates the result of a reduction of this conjunction by at least one conjunct. Moreover, the second and third sufficient conditions $U\overline{L}$ and $\overline{U}L$ contain one sufficient proper part each: U and L respectively. As both U and L do not contain any further proper parts, they are not further minimizable. Thus, U and L each are minimally sufficient for G—or in the terminology of Quine-McCluskey optimization: U and L are the two *prime implicants* of G.

The feature of this minimization procedure that will be of crucial importance for the sequel of this paper is that Quine-McCluskey optimization only eliminates conjuncts of a sufficient condition if the input table actually contains a pair of rows that accord with respect to the outcome as well all (atomic) conditions except for one. If such a pair of rows does not exist for a particular sufficient condition, the latter cannot be further minimized. To facilitate later reference to this restriction, I furnish it with a label: I shall speak of the *one-difference restriction*.

Before we look at the consequences of the one-difference restriction, let us conclude this overview of the basics of QCA. After minimizing sufficient conditions, QCA first identifies and then minimizes necessary conditions of the outcome. In case of table 1, this final part of QCA is straightforward. Every country considered in our hypothetical study that provides a generous welfare state also has either strong unions or strong left parties. That is, the disjunction $U \vee L$ is necessary for G. Moreover, $U \vee L$ does not contain necessary proper parts, where a *proper part of a disjunction* $Z_1 \vee Z_2 \ldots \vee Z_n$ designates the result of any reduction of this disjunction by at least one disjunct. In our example, neither U nor L are themselves necessary for G, for there are cases in which G is given without U and cases featuring G without L. $U \vee L$ is hence *minimally* necessary for G.

Depending on investigated research questions, QCA can then be reapplied to identify and minimize sufficient and necessary conditions for the absence of the outcome. c_4 is the only row in table 1 featuring $\overline{G}$. The configuration of conditions in c_4 is sufficient for $\overline{G}$: $\overline{UL} \rightarrow \overline{G}$. As there is no other row satisfying the one-difference restriction with respect to c_4, $\overline{UL}$ cannot be further minimized. $\overline{UL}$ is, hence, minimally sufficient for $\overline{G}$. Moreover, $\overline{UL}$ accounts for all occurrences of $\overline{G}$: $\overline{G} \rightarrow \overline{UL}$. Since $\overline{UL}$ does not contain any necessary proper parts, it is minimally necessary for $\overline{G}$. At the end, QCA formally integrates all the uncovered relations of minimal sufficiency and necessity in so-called *solution formulas*. Our exemplary QCA analysis produces the following solution formulas:

$$U \vee L \leftrightarrow G \quad ; \quad \overline{UL} \leftrightarrow \overline{G} \tag{1}$$

Finally, as all redundancies are removed from its solutions formulas, QCA proceeds to causally interpret the dependencies expressed in these formulas. In our example, QCA rules that strong unions and strong left parties are alternative causes of welfare generosity and that their joint absence is a complex cause of a stingy welfare system.

Of course, the identification of minimally sufficient and necessary conditions is not normally as straightforward as in case of table 1. One problem that regularly affects the analysis of small-N data is of particular relevance for our current purposes, because, in combination with the one-difference restriction, this problem imposes considerable constraints on the causal interpretability of corresponding data and on the causal complexity uncovered by QCA. As Ragin and Sonnett (2005, 180) put it:

> Naturally occurring social phenomena are limited in their diversity. In fact, it could be argued that limited diversity is one of their trademark features.

In the terminology of QCA, the diversity of configurational data is said to be *limited* if not all 2^n configurations of n conditions of an investigated outcome are contained in these data (cf. Ragin 2000, 139). Such limitation may occur for a host of different reasons. Social scientists are inevitably confined to the variety of cases social reality and history happen to provide for them.

To make the problem arising from limited diversity more concrete, consider table 2 which lists types of countries of another hypothetical study of the causal connections between U, L, and G. Table 2 does not feature all 2^2 logically possible configurations of the two conditions U and L. Supposedly, in our second study we did not find countries with strong unions and weak left parties, i.e. cases of type $U\overline{L}$ are absent from the data. Such a missing configuration is commonly termed

#	U	L	G
c_1	1	1	1
c_2	0	1	1
c_3	0	0	0

Table 2. Second hypothetical data table listing configurations of U (strong unions), L (strong left parties), and G (generous welfare state).

a *logical remainder*. There are two sufficient conditions for G in table 2, *viz.* the conjunction of strong unions and strong left parties (UL in c_1) and the conjunction of weak unions and strong left parties ($\overline{U}L$ in c_2). Furthermore, the pair of rows $\langle c_1, c_2 \rangle$ satisfies the one-difference restriction and establishes that L alone is sufficient for G, i.e. that L is minimally sufficient for G. However, the question remains whether U is also minimally sufficient for G (or whether L is moreover necessary for G). To answer that question we would have to know whether countries with strong unions and weak left parties provide generous welfare systems or not. But since the data available to us do not exhibit any countries of type $U\overline{L}$, it is empirically undetermined what value G would take in cases of this type. Furthermore, table 2 features one sufficient condition for $\overline{G}$, *viz.* the configuration $\overline{UL}$ in row c_3. As no other row accords with c_3 in regard to all conditions except for one, $\overline{UL}$ cannot be further minimized. However, if cases of type $U\overline{L}$ were in fact to exhibit $\overline{G}$, $\overline{UL}$ would be minimizable, for it would then turn out that $\overline{L}$ is itself sufficient for $\overline{G}$. These ambiguities illustrate the problem of limited diversity: since no case of type $U\overline{L}$ is contained in the data, it is undeterminable, from the perspective of QCA, whether both U and L are causes of G or G is only caused by L and whether $\overline{UL}$ is a complex cause of $\overline{G}$ or not.

Empirical indeterminacies of this type can only be resolved if prior theoretical knowledge is available about the causal dependencies among investigated conditions and outcomes. Such a theoretical background may have different implications on whether logical remainders could possibly have been instantiated in analyzed cases or not and on what values the outcomes would have taken, had remainders in fact been observed. That means background theories may have different ramifications for *counterfactual cases*. To do justice to these differences in background knowledge, Ragin and Sonnett (2005) distinguish three different strategies researchers may adopt when analyzing limitedly diverse data. According to the first and most conservative strategy—call it $\mathcal{S}_1$—, remainders are taken to be excluded (or false), i.e. relevant background knowledge tells the researcher that corresponding remainders could under no circumstances have been observed. As to the second, intermediate strategy—$\mathcal{S}_2$—, remainders are determined to be empirically possible by background knowledge, which moreover supplies enough information to decide which values an investigated outcome would have taken, had a pertaining remainder in fact been observed. Finally, the third and most liberal strategy—$\mathcal{S}_3$—treats remainders as so-called *don't care* cases, i.e. as empirically

#	U	L	G
c_1	1	1	1
c_2	0	1	1
c_3	0	0	0
c_4^*	1	0	1

(a)

#	U	L	G
c_1	1	1	1
c_2	0	1	1
c_3	0	0	0
c_4^{**}	1	0	0

(b)

Table 3. The two possible counterfactual completions of table 2.

possible cases for which outcomes may be set to whichever value yields the most parsimonious solution formulas. In the terminology of QCA, *don't care* cases are said to be available as *simplifying assumptions*.

While the details of these strategies, which, among other things, involve intricate assessments of how "easy" or "difficult" relevant counterfactual claims are, are of no concern to us here, it is important to note that the strategies generate different solution formulas. I illustrate these differences by means of the hypothetical study in table 2. $\mathcal{S}_1$ does not add counterfactual cases to table 2. In consequence, the question whether U is also minimally sufficient for G or L is moreover necessary for G has to be left open. Moreover, as the one-difference restriction is not met for the one sufficient condition of $\overline{G}$, *viz.* for $\overline{UL}$, it cannot be further minimized. All in all, $\mathcal{S}_1$ produces the following solution formulas for G and $\overline{G}$, respectively:

$$L \rightarrow G \quad ; \quad \overline{UL} \rightarrow \overline{G} \tag{2}$$

Contrary to $\mathcal{S}_1$, both $\mathcal{S}_2$ and $\mathcal{S}_3$ introduce the remainder $U\overline{L}$ as a counterfactual case. Depending on what outcome value is assigned to this case, the completion of table 2 yields either table 3a or 3b, where c_4^* and c_4^{**} designate the two conceivable counterfactual cases. According to $\mathcal{S}_2$, the value of G is to be determined by the researcher's theoretical background. Let us assume that our currently best background theories on welfare systems entail that if a country had strong unions and weak left parties, it would provide a generous welfare state. That is, we complete table 2 by the counterfactual case c_4^* and obtain table 3a. In this table, L is not necessary for G, as welfare generosity may also occur without strong left parties, namely in countries with strong unions. G has two minimally sufficient conditions in table 3a, U and L, whose disjunctive concatenation, $U \vee L$, is minimally necessary for G. Moreover, as table 3a does not satisfy the one-difference restriction for $\overline{G}$'s sufficient condition $\overline{UL}$, the latter turns out not to be further minimizable. $\overline{G}$ occurs if and only if $\overline{UL}$ occurs. In sum, $\mathcal{S}_2$ produces the following solution formulas relative to 3a:

$$U \vee L \leftrightarrow G \quad ; \quad \overline{UL} \leftrightarrow \overline{G} \tag{3}$$

Finally, $\mathcal{S}_3$ also supplements the counterfactual configuration $U\overline{L}$ to generate either table 3a or 3b. Contrary to $\mathcal{S}_2$, however, $\mathcal{S}_3$ does not make the choice between 3a or 3b dependent on background theories, but simply chooses the table that produces the more parsimonious solution formulas. In our exemplary case, parsimony is maximized if countries with strong unions and weak left parties are assumed not to provide a generous welfare state, i.e. if we settle for table 3b. In this table, the one-difference restriction is satisfied for the sufficient conditions of both G and $\overline{G}$. There is a row that accords with c_1 in all but one respect, *viz.* c_2, and that allows for the elimination of U from the sufficient condition UL of G; likewise, there is a row that accords with c_3 in all but one respect, *viz.* c_4^{**}, and that allows for the elimination of $\overline{U}$ from the sufficient condition $\overline{UL}$ of $\overline{G}$. Furthermore, L and $\overline{L}$ are each not only sufficient but also necessary for G and $\overline{G}$, respectively. Overall, $\mathcal{S}_3$ yields the following solution formulas:

$$L \leftrightarrow G \quad ; \quad \overline{L} \leftrightarrow \overline{G} \tag{4}$$

Plainly, tables 2 and 3 constitute very simple examples. Relative to more complex data tables, differences between solution formulas produced by $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$ tend to be far greater. However, the solution formulas for our simple examples have one commonality which they share with all QCA solution formulas, independently of the data's complexity: QCA always directly connects conditions to outcomes. Depending on minimization strategies chosen, solution formulas may differ with respect to the complexity of identified complex or alternative causes, but QCA only assigns *direct causes* to outcomes. More concretely, QCA either determines strong unions and strong left parties to be alternative causes of the generosity of the welfare system or strong left parties are identified as a both sufficient and necessary cause of a generous welfare state. Under no circumstance would QCA conclude that conditions U and L are themselves causally connected. QCA never models input data in terms of causal chains.

This is a direct consequence of minimizing sufficient conditions on the basis of Quine-McCluskey optimization, which imposes the one-difference restriction. Even though our initial input table 2 features no case such that strong unions are combined with weak left parties, QCA requires the introduction of such a remainder as a counterfactual case in order to assess the minimality of sufficiency and necessity relationships. Whenever an input table does not contain 2^n configurations of n conditions, QCA takes that table to be limited in its diversity. However, that may be a hasty conclusion. It might well be that table 2 in fact contains *all* empirically possible configurations of strong unions and strong left parties, because these two conditions themselves might be *causally dependent*. As to table 2, every country with strong unions also has strong left parties, i.e. U is sufficient for L. This dependency must by no means stem from historical contingencies but could be the result of U being a cause of L. That is, the data in table 2 might result from a causal chain such that U is a cause of L which is a cause of G. Row c_2, that features L without U, moreover indicates that U cannot be the only cause of L, for U is not necessary for L. Accordingly, there exists at least one (unknown) alternative cause Z of L which is not among the conditions considered in table 2. Overall, the data in table 2 might stem from a causal structure as depicted in figure 1.

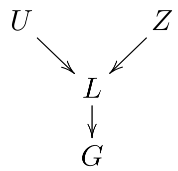

Figure 1. A causal chain model that fits the data in table 2.

It is beyond doubt that many social phenomena result from causal chains (cf. Goertz 2006). In fact, chances are that the strength of unions and the strength of left parties—at least in democratic countries—are tightly causally dependent. In that case, both strategies $\mathcal{S}_2$ and $\mathcal{S}_3$, by introducing the remainder $U\overline{L}$ as a counterfactual case, distort the data that they intend to model in a way that *violates* the actually underlying causal structure. The remainder $U\overline{L}$ is not compatible with U being a sufficient cause of L. Hence, if the data in table 2 is really the result of the chain in figure 1, both $\mathcal{S}_2$ and $\mathcal{S}_3$ generate causal models that severely misrepresent the actual causal structure. By abstaining from introducing counterfactual cases, strategy $\mathcal{S}_1$ does not fallaciously distort the data, if table 2 indeed stems from a chain. Nonetheless, $\mathcal{S}_1$ produces inadequate solution formulas. $\mathcal{S}_1$ does not properly minimize the sufficiency relationships in table 2. Just as $\mathcal{S}_2$ and $\mathcal{S}_3$, $\mathcal{S}_1$ fails to recognize the sufficiency of U for L and G and, thus, the direct causal relevance of strong unions for strong left parties and the indirect relevance of strong unions for welfare generosity. In sum, notwithstanding the fact that the chain in figure 1 perfectly fits the data in table 2, none of the search strategies supplied by QCA succeeds in modeling table 2 in terms of that chain.

Although the literature on QCA currently does not provide any other strategies to process limitedly diverse data (cf. also Schneider and Wagemann 2010, 408), it might be argued that QCA could be amended by a further search strategy that assigns the chain in figure 1 to table 2 after all. In particular, it might be held that a subdivision of causal chains into their separate layers yields causal substructures that are amenable to a *stepwise* QCA analysis. Indeed, Schneider and Wagemann (2006) have suggested a stepwise application of QCA to remote and proximate conditions of an outcome in order to distinguish among relevant background contexts in which proximate conditions are causally efficacious. Even though this so-called *Two-Step* approach is not designed to uncover causal chains, something along its lines might be proposed as a new QCA strategy to process chain-generated data. Such a search strategy—call it $\mathcal{S}_4$, for short—could be roughly spelled out as follows: in a first phase, QCA is applied to identify the direct causes of the ultimate outcome among the conditions; in a second phase, QCA is sequentially reapplied to uncover the causal dependencies among the conditions themselves.

Let us investigate whether $\mathcal{S}_4$ could indeed model table 2 in terms of the causal chain in figure 1. In the first phase of $\mathcal{S}_4$, we hence apply QCA to identify the direct causes of G in the set $\{U, L\}$. Here, the limited diversity of table 2 again raises the question how to handle logical remainders. If no remainders are counterfactually added, as in case of strategy $\mathcal{S}_1$, QCA is not able to determine whether L is the only direct cause of G or whether U is also directly relevant to G. In order to infer that L is the only direct cause of G, as in the chain of figure 1, QCA has to treat the logical remainder $U\overline{L}$ as *don't care* case in the first phase of $\mathcal{S}_4$—analogously to $\mathcal{S}_3$. Such a coding of $U\overline{L}$ makes the case c_4^{**} available as simplifying assumption. However, by counterfactually supplying the case c_4^{**} to yield table 3b, a simplifying assumption is introduced that not only prompts QCA to infer that L is the only direct cause of G but also (in combination with the other rows of table 3b) entails that U and L are *independent*. This independence, in turn, contradicts the causal chain in figure 1. As a consequence, in its first phase, strategy $\mathcal{S}_4$ inevitably faces a dilemma: either it cannot establish L as only direct cause of G or it is forced to counterfactually add a logical remainder that renders U and L independent and, hence, violates the causal chain that is being searched. Neither horn of that dilemma results in an analysis of table 2 in terms of the chain in figure 1. Choosing the first horn reduces strategy $\mathcal{S}_4$ to $\mathcal{S}_1$, whereas choosing the second horn reduces strategy $\mathcal{S}_4$ to $\mathcal{S}_3$.

These considerations suggest that the problems QCA faces when confronted with chain-generated data do not stem from the current (accidental) unavailability of a proper search strategy for such data. Rather, these problems stem from the calculative core of the method. Minimizing configurational data on the basis of Quine-McCluskey optimization presupposes that all 2^n logically possible combinations of n analyzed conditions are empirically possible, which, in turn, presupposes that there are no causal dependencies among those n conditions. That is, by resorting to Quine-McCluskey optimization and, hence, by subscribing to the one-difference restriction, the QCA framework assumes that analyzed conditions are mutually causally independent. For later reference I label this the *independence assumption*, or (IND) for short. Even though, to my knowledge, Ragin has never explicitly stated that QCA is only a correct method if (IND) is assumed, he tailors his notion of *causal complexity* to the limitations (IND) imposes 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). In fact, if causal structures underlying configurational data are assumed to have a maximal complexity as defined in this quotation, (IND) is satisfied. I take this to indicate that QCA is, from the outset, designed to analyze causal structures featuring exactly *one* effect and a possibly complex configuration of *mutually independent direct* causes of that effect.[2]

Apart from making explicit these limitations on the causal complexity which is correctly discoverable by QCA, these considerations show that a methodology of configurational causal reasoning that correctly models data stemming from chains must avoid the one-difference restriction. An alternative methodology which does not impose that restriction has been introduced in Baumgartner (2009a,b). It has been termed *Coincidence Analysis*, or CNA for short. The next section reviews the basic idea behind CNA's alternative minimization procedure.

3. THE BASICS OF COINCIDENCE ANALYSIS

Coincidence Analysis shares all of QCA's basic goals and intentions. It focuses on configurational complexity rather than on net effects, it processes the same kind of data as QCA, and it implements the same regularity theoretic notion of causation, as e.g. developed by Mackie (1974). Apart from its altered minimization procedure for sufficient and necessary conditions which will be presented below, there is one difference between QCA and CNA that deserves separate mention at this point. Contrary to QCA, CNA does not presuppose that analyzed variables can be classified into potential causes and a corresponding outcome prior to analyzing the data. If such a classification is available, so much the better; if not, CNA simply identifies and minimizes all relationships of sufficiency and necessity that subsist among the relevant variables and issues a set of causal models that all entail these sufficiency and necessity relations. It is then up to the researcher and her background theories to choose among these possible models.

Accordingly, CNA does not normally distinguish between conditions and outcomes, rather it just speaks neutrally of *factors*. 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*. Moreover, to reflect the fact that causally interacting factors are co-instantiated within the same spatiotemporal region, i.e. coincidently, configurations of analyzed factors are termed *coincidences* in the CNA-context rather than *cases*—which explains the name "Coincidence Analysis". All of these are mere terminological differences. Nothing substantial hinges on them. Accordingly, instead of "Coincidence Analysis" one might just as well speak of "Case Analysis"—or even of "causal-chain-QCA" ($ccQCA$) for that matter; for, as will be shown in the remainder of this paper, the one substantial difference between CNA and QCA is that, contrary to the latter, the former can correctly process data tables that violate (IND). More specifically, contrary to QCA, CNA does not minimize relationships of sufficiency and necessity by means of Quine-McCluskey optimization, but based on its own custom-built minimization procedure.

The basic idea behind this procedure can be easily stated. If there exists any kind of (deterministic) causal dependency among n factors, it follows that not all 2^n logically possible configurations of these factors are also empirically possible. Causal dependencies constrain the range of empirical possibilities. To do justice to this trademark feature of causality, CNA does not only infer causal dependencies from the coincidences (or cases) actually contained in data tables, but also from the coincidences *not contained therein*. In fact, evidence as to empirically impossible coincidences is of central relevance for causal discovery. Claims about sufficiency and necessity are logically equivalent to negative existential claims. For example, to state that strong left parties are sufficient for welfare generosity is equivalent to stating that there are no cases featuring strong left parties and a weak welfare state. Analogously, claiming that strong left parties are necessary for welfare generosity is equivalent to claiming that there are no cases featuring welfare generosity without strong left parties. Negative existentials of this sort constitute the core of CNA's minimization procedure: to determine whether, say, a complex sufficient condition $Z_1 Z_2 \ldots Z_m$ of a factor Z_n contains redundancies or not, CNA parses a corresponding data table to check whether the table contains a row featuring a proper part of that sufficient condition, say $Z_2 \ldots Z_m$, in combination with $\overline{Z_n}$ or not. If the table does not contain such a such row, $Z_2 \ldots Z_m$ is itself sufficient for Z_n, i.e. Z_1 is redundant. Next, $Z_2 \ldots Z_m$ is likewise tested for further redundancies, and so forth, until no more redundancies are found—and analogously for necessary conditions.

To make all of this more precise, some notational and terminological preliminaries are required. Factors are symbolized by italicized capital letters A, B, C, etc., with variables (placeholders) Z, Z_1, Z_2, etc. running over the domain of factors. The negation of a factor A is written as before: $\overline{A}$. Moreover, I introduce variables X_1, X_2, etc. that run over the domain of coincidences (configurations) of an open number of factors. Causal analyses are always relativized to a set of investigated factors. To this set I refer as the *factor frame* of the analysis. As indicated above, CNA does not presuppose that a particular factor from the frame can be identified as the outcome of an analyzed causal structure prior to applying CNA. CNA simply identifies all relationships of sufficiency and necessity among the factors in the frame and properly minimizes these relationships. In sociological practice, however, it is often known from the outset which factors are possible causes and which ones are possible effects. Accordingly, in addition to a data table, CNA may be given a subset W of possible effects from the frame as input. Sufficient and necessary conditions are then calculated for the members of W only.

CNA then first identifies minimally sufficient conditions for each of the factors in W. This is done in four steps: (i) a factor $Z_i \in \mathsf{W}$ is selected, (ii) all sufficient conditions of Z_i are identified, (iii) these sufficient conditions are minimized, and (iv) the procedure is restarted at (i) by selecting another $Z_j \in \mathsf{W}$, until all factors in W have been selected. By referring to the other factors in the frame apart from a selected Z_i as *residuals*, the rule that identifies sufficient conditions of Z_i in a given input table $\mathcal{C}$ can be stated as follows:

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

A complex sufficient condition X_k of Z_i contains no redundancies if and only if X_k contains no sufficient proper parts, i.e. if no elimination of a conjunct of X_k results in a condition that is itself sufficient for Z_i. More precisely put:

(MSUF) A sufficient condition $Z_1 Z_2 \ldots Z_h$ of Z_i is *minimally* sufficient if and only if neither $Z_2 Z_3 \ldots Z_h$ nor $Z_1 Z_3 \ldots Z_h$ nor ... nor $Z_1 Z_2 \ldots Z_{h-1}$ are sufficient for Z_i according to (SUF).

To test whether a sufficient condition X_k of Z_i is minimally sufficient in the sense defined by (MSUF), every factor in X_k is to be tested for redundancy by eliminating it from that condition and checking whether the remaining condition still is sufficient for Z_i. A sufficient condition of Z_i is minimally sufficient if and only if every elimination of a factor from that condition results in the insufficiency of the remaining condition. This can be more formally put as follows:

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

The core difference between minimizing sufficient conditions along the lines of Quine-McCluskey optimization and of (MSUF') deserves separate emphasis: Quine-McCluskey optimization only eliminates conjuncts of a sufficient condition if the latter reduced by a respective conjunct is actually contained in the data table in a way that satisfies the one-difference restriction; by contrast, (MSUF') eliminates conjuncts of a sufficient condition if the latter reduced by a respective conjunct is not contained in the data in combination with the absence of a corresponding effect.

To illustrate CNA's minimization of sufficient conditions, reconsider table 1. For simplicity, assume that our theoretical background—as in case of the exemplary QCA analysis of table 1 conducted in section 2—determines G to be the only conceivable effect among the three factors contained in that table, i.e. $\mathsf{W} = \{G\}$. Rows c_1, c_2, and c_3 each feature a sufficient condition of G according to (SUF): For UL, $U\overline{L}$, and $\overline{U}L$ table 1 contains one row featuring ULG, $U\overline{L}G$, and $\overline{U}LG$, respectively, and no row in which those conditions are combined with $\overline{G}$. Moreover, no row in table 1 exhibits either U or L in combination with $\overline{G}$. That is, both U and L are themselves sufficient for G. As neither of them has further proper parts, U and L are each minimally sufficient for G. Analogous considerations reveal that $\overline{UL}$ is minimally sufficient for $\overline{G}$ in table 1. Row c_4 exhibits the coincidence $\overline{ULG}$ and for each proper part of $\overline{UL}$ there is a row featuring that part in combination with the absence of $\overline{G}$, *viz.* $\overline{U}G$ in c_3 and $\overline{L}G$ in c_2. Accordingly, none of the proper parts of $\overline{UL}$ is itself sufficient for $\overline{G}$.

Next, CNA disjunctively combines minimally sufficient conditions of each $Z_i \in \mathsf{W}$ to necessary conditions of Z_i. Necessity of a disjunction of conditions relative to a given input table $\mathcal{C}$ is defined as follows:

(NEC) A disjunction $X_1 \vee X_2 \vee \dots \vee X_h$ of minimally sufficient conditions of Z_i is *necessary* for Z_i if and only if $\mathcal{C}$ contains no row featuring Z_i in combination with $\neg(X_1 \vee X_2 \vee \dots \vee X_h)$, i.e. no row comprising $\overline{X_1X_2\dots X_h}Z_i$.

Finally, if CNA finds necessary conditions of $Z_i \in \mathsf{W}$, it proceeds to minimize those conditions analogously to (MSUF) and (MSUF').

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

To determine whether a necessary condition $X_1 \vee X_2 \vee \ldots \vee X_h$ of Z_i is minimally necessary in the sense defined by (MNEC), every disjunct contained in $X_1 \vee X_2 \vee \ldots \vee X_h$ is to be tested for redundancy by eliminating it from that disjunction and checking whether the remaining condition still is necessary for Z_i. A necessary condition of Z_i is minimally necessary if and only if every elimination of a disjunct results in the loss of necessity of the remaining condition. More formally and 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.

To illustrate, let us again apply these rules to table 1. Above, we saw that U and L are each minimally sufficient for G and that $\overline{UL}$ is minimally sufficient for $\overline{G}$. As it turns out, the disjunctive concatenation of U and L, *viz.* $U \vee L$, accounts for all occurrences of G in table 1. That is, $U \vee L$ is necessary for G. Analogously, $\overline{UL}$ accounts for all occurrences of $\overline{G}$ in this table, i.e. $\overline{UL}$ is necessary for $\overline{G}$. Moreover, neither U nor L are themselves necessary for G, because for both of them there is a row where they are absent while G is given: c_3 features the coincidence $\overline{U}G$ and c_2 the coincidence $\overline{L}G$. Therefore, $U \vee L$ is minimally necessary for G. Finally, as $\overline{UL}$ has no necessary proper parts either, it is minimally necessary for $\overline{G}$. All in all, CNA produces the following solution formulas for table 1:

$$U \vee L \leftrightarrow G \quad ; \quad \overline{UL} \leftrightarrow \overline{G} \tag{5}$$

It can easily be seen that this solution is the same as the solution assigned to table 1 by QCA, i.e. (1). While QCA only eliminates redundant elements from sufficient and necessary conditions if the one-difference restriction is satisfied, CNA systematically tests for eliminability, independently of whether the one-difference restriction is satisfied or not. Yet, evidently, if a data table features all 2^n configurations of n residuals of some $Z_i \in \mathsf{W}$, as does table 1 for the residuals of G, QCA can perform the same systematic redundancy testing which CNA performs independently of the availability of all those logically possible configurations. Consequently, if the data exhibit all 2^n configurations of n residuals, CNA and QCA produce the exact same solution formulas. The two methodologies are equivalent for all data tables that are logically complete in this sense. As the next section is going to show, though, important differences emerge if input tables are not logically complete with respect to residuals of some $Z_i \in \mathsf{W}$ and the theoretical background of the researcher has it that there might exist causal dependencies among residuals.

4. CNA AND CAUSAL CHAINS

In order to have a concrete background against which to discuss how CNA processes data that are generated by causal chains, let us return to table 2. Assume, a hypothetical study on the causal dependencies among strong unions (U), strong left parties (L), and welfare generosity (G) has generated table 2; and assume furthermore that this data in fact is the result of the causal chain depicted in figure 1. In consequence, table 2 lists *all empirically possible combinations* of the three factors U, L, and G. Or differently, table 2 is not limited in its diversity, even though it does not contain coincidences featuring strong unions and weak left parties. Adding counterfactual coincidences to this table would hence violate the underlying causal structure.

As we have seen in section 2, the only QCA-strategy that does not add counterfactual cases to table 2, strategy $\mathcal{S}_1$, does not succeed in recognizing the dependencies between U and L and between U and G as being of causal nature. Let us now apply CNA to that table. Available prior causal knowledge yields that welfare generosity must be the ultimate outcome of the causal structure we are looking for. We have enough evidence indicating that countries install generous welfare systems only (temporally) *after* unions or left parties have gained sufficient strength. However, suppose we have no theoretical knowledge about the causal interplay between the strength of unions and the strength of left parties. On the face of it, there may exist any kind of causal relationship between these two factors. In light of this, CNA is brought to bear in such a way that, first, it identifies sufficient and necessary conditions for all factors in the frame and, second, the researcher selects those relationships of sufficiency and necessity that can be causally interpreted relative to the available theoretical background.

We thus start by setting the set W of potential effects equal to the factor frame, i.e. $\mathsf{W} = \{U, L, G\}$, such that CNA identifies minimally sufficient and necessary conditions for all factors in the frame. For simplicity, we abstain from also identifying sufficiency and necessity relations for the absences of U, L, and G. Rows c_1 and c_2 of table 2 each feature a sufficient condition of G according to (SUF): For both UL and $\overline{U}L$ the table contains one row featuring ULG and $\overline{U}LG$, respectively, and no row in which those conditions are combined with $\overline{G}$. Moreover, no row in table 2 exhibits either U or L in combination with $\overline{G}$. That is, both U and L are themselves sufficient for G. As neither of them has proper parts, U and L are each minimally sufficient for G according to (MSUF). Very analogous considerations reveal that U and G are each minimally sufficient for L in table 2. However, as instances of G generally occur temporally after instances of L, a causal interpretation of the sufficiency of G for L can be excluded to begin with. Finally, there are no sufficient conditions of U in table 2. The coincidence LG is combined with U in c_1 and with $\overline{U}$ in c_2, and is therefore not sufficient for U according to (SUF). Overall, the first stage of our CNA-analysis of table 2 yields the following causally interpretable minimally sufficient conditions of the members of W:

$$\begin{aligned} G &: \{U, L\} \\ L &: \{U\} \\ U &: \{\} \end{aligned}$$

CNA then proceeds to build necessary conditions of the members of W from this inventory of their minimally sufficient conditions. The disjunction $U \vee L$ is necessary for G according to (NEC), for there is no row in table 2 exhibiting the coincidence $\overline{U}\overline{L}G$. Moreover, $U \vee L$ has a necessary proper part: There is no row featuring $\overline{L}$ in combination with G. That is, as to (MNEC) L is minimally necessary for G. By contrast, the one minimally sufficient condition of L that is amenable to a causal interpretation, i.e. U, is not necessary for L. In row c_2, L is instantiated without U. That shows that L has causes that are not contained in the frame of our exemplary study. As there are no minimally sufficient conditions of U in table 2, CNA does not build any necessary conditions of U either.[3] Overall, CNA hence finds one causally interpretable minimally necessary condition composed of minimally sufficient conditions for G, i.e. L, and one causally interpretable minimally sufficient condition of L, i.e. U. In the final solution formula, CNA conjunctively concatenates its findings:

$$(U \rightarrow L) \wedge (L \leftrightarrow G) \tag{6}$$

(6) mirrors exactly the sufficiency and necessity relations that follow from the causal chain in figure 1. Moreover, (6) generates exactly the truth table 2, that is, (6) is true if and only if U, L, and G are assigned one of the configurations of truth values listed in table 2. Contrary to QCA, CNA does not eliminate the dependency between U and L, but recognizes it as being of causal nature. By allowing for more than one outcome and by systematically testing sufficient and necessary conditions for redundancies, independently of whether the one-difference restriction is satisfied, CNA succeeds in adequately modeling the data in table 2 in terms of the causal chain in figure 1.

It must be emphasized that CNA's assignment of (6) to table 2 essentially hinges on the presumed *empirical completeness* of that table. If a researcher is simply confronted with a table as 2, she cannot be sure that the absence of countries with strong unions and weak left parties from the data is due to a causal dependence between these factors. The fact that this configuration is missing from the data might also be the result of an accidental limitation of data diversity. Plainly, configurational data themselves do not shed any light whatsoever on why certain configurations are not contained therein. Explanations for missing data points must come from external sources—in the first instance, from the researcher's theoretical background. If available background knowledge about the interplay between the strength of unions and of left parties suggests that the combination of strong unions and weak left parties is empirically possible after all, a CNA-analysis has to revert to counterfactual completions of data tables along the lines of strategies $\mathcal{S}_2$ and $\mathcal{S}_3$ sketched in section 2. As we have seen in section 3, since CNA is equivalent to QCA with respect to logically complete tables, CNA generates the same solution formulas as strategies $\mathcal{S}_2$ and $\mathcal{S}_3$, depending on whether table 2 is counterfactually completed in terms of tables 3a or 3b.

While in experimental disciplines researchers could freely manipulate investigated factors and, thus, test whether missing configurations can be artificially produced or not, assessing the empirical completeness of configurational data may give rise to severe problems in non-experimental disciplines—for example, if the theoretical background is indeterminate as to whether the combination of strong unions and weak left parties is compatible with the underlying causal structure or not. Deciding whether missing configurations are due to accidental diversity limitations or to causal dependencies constitutes one of the trademark problems of causal reasoning in non-experimental disciplines. Nonetheless, this *problem of data completeness* has received far less attention in the literature on configurational causal reasoning than the problem of limited diversity.

I have deliberately chosen very simple data tables here to focus on the computational differences between QCA and CNA. In more complex cases, of course, data tables typically feature a host of logical remainders. Realistically, only a subset of all remainders will stem from causal dependencies among residuals. One of the central upshots of this paper is that supplementing counterfactual cases along the lines of strategies $\mathcal{S}_2$ and $\mathcal{S}_3$ is only warranted for those factors that are determined to be independent by the theoretical background, i.e. for those configurations of residuals that are determined to be empirically possible. The first step in the causal analysis of small- and intermediate-N data always has to be to consult the available theoretical background in order to counterfactually supplement those conditions that are missing from the data due to mere contingencies of data collection. Procedures of Boolean causal reasoning can only be brought to bear *after an exhaustive* collection of data. In this regard, the computational difference between QCA and CNA entails another important difference between the two methods: CNA is more liberal than QCA with respect to what can count as an exhaustive collection of configurational data. While QCA requires 2^n combinations of n conditions for proper minimizations of solution formulas, CNA can properly minimize any number of combinations smaller than 2^n. CNA can process data that stem from causal structures involving both *multiple* effects and *mutually dependent* causes. Contrary to QCA, CNA does not need to assume (IND). If (IND) does not hold, sufficient and necessary conditions must not be minimized along the lines of Quine-McCluskey optimization, but along the lines of (MSUF') and (MNEC').

Generally, configurational data greatly underdetermine their own causal analysis. Such data do not themselves distinguish between conditions and outcomes and they do not wear their own completeness on their sleeves. In non-experimental disciplines, the distinction between conditions and outcomes or the completeness of relevant data must be determined by non-configurational information, most notably, by the available theoretical background. Without such additional information, the causal analysis of configurational data is inevitably ambiguous. The more theoretical guidance is available, the more modeling ambiguities can be resolved. As indicated in section 3, CNA aims to make all causal models explicit that fit the data relative to any given theoretical background. While QCA can only be applied if a single factor is identifiable as outcome and the remaining conditions are determined to be causally independent, CNA is applicable even without any such theoretical guidance. Clearly though, without such guidance the set of possible models assigned to a data table by CNA will commonly be rather large. But that is just what we should expect from a method of causal inference: if the data and the background theories underdetermine causal modeling, an inference procedure must bring all data-fitting models on the table, independently of the number of the resulting model candidates.

Notes

[1]Instead of this logical or truth-functional terminology the QCA literature often draws on a set-theoretical vocabulary. That is, instead of *sufficient conditions* QCA is said to detect *subset relations*, or instead of the *conjunction* of two conditions authors talk of the *intersection* of two sets. These two terminologies are entirely equivalent. For mere reasons of taste I consistently use the logical terminology in this paper. Explicit translations between the two terminologies can be found in Goertz (2003).

[2]Caren and Panofsky (2005) have generalized QCA for temporally ordered conditions ($TQCA$). Even though Caren and Panofsky are not very explicit about whether they take temporal order among conditions to indicate causal dependencies, $TQCA$ clearly still relies on Quine-McCluskey optimization (cf. Caren and Panofsky 2005, 156) and, hence, yields incorrect solution formulas for chain-generated data.

[3]Even though both L and G are necessary for U in table 2, CNA does not causally interpret these dependencies. While causes are sometimes necessary for their effects, effects are always necessary for their causes, for when no effects occur, no causes occur either. Therefore, if a factor Z_1 is necessary for another factor Z_2, additional evidence is needed to establish that the direction of causation is from Z_1 to Z_2 and not the other way around. While QCA simply presupposes a causal order, CNA identifies the direction of causation via relationships of sufficiency. Without corresponding sufficiency relationships CNA abstains from causally interpreting relationships of necessity. For details cf. Baumgartner (2009a).

REFERENCES

Baumgartner, M. (2008), "Regularity Theories Reassessed," *Philosophia*, 36, 327–354.

——— (2009a), "Inferring Causal Complexity," *Sociological Methods & Research*, 38, 71–101.

——— (2009b), "Uncovering Deterministic Causal Structures: A Boolean Approach," *Synthese*, 170, 71–96.

Caren, N. and Panofsky, A. (2005), "TQCA: A Technique for Adding Temporality to Qualitative Comparative Analysis," *Sociological Methods & Research*, 34, 147–172.

Goertz, G. (2003), "Assessing the Importance of Necessary or Sufficient Conditions in Fuzzy-Set Social Science," Technical report, Compasss Working Paper WP2003-7.

——— (2006), *Social Science Concepts: A User's Guide*, Princeton: Princeton University Press.

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

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. and Sonnett, J. (2005), "Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases, and Comparative Analysis," in *Vergleichen in der Politkwissenschaft*, VS Verlag für Sozialwissenschaften.

Schneider, C. Q. and Wagemann, C. (2006), "Reducing Complexity in Qualitative Comparative Analysis (QCA): Remote and Proximate Factors and the Consolidation of Democracy," *European Journal of Political Research*, 45, 751–786.

——— (2010), "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets," *Comparative Sociology*, 9, 397–418.