THE CAUSAL CHAIN PROBLEM

Michael Baumgartner

Abstract

This paper addresses a problem that arises when it comes to inferring deterministic causal chains from pertaining empirical data. It will be shown that to every deterministic chain there exists an empirically equivalent common cause structure. Thus, our overall conviction that deterministic chains are one of the most ubiquitous (macroscopic) causal structures is underdetermined by empirical data. It will be argued that even though the chain and its associated common cause model are empirically equivalent there exists an important asymmetry between the two models with respect to model expansions. This asymmetry might constitute a basis on which to disambiguate corresponding causal inferences on non-empirical grounds.

1 Introduction

No later than Frydenberg (1990), Verma and Pearl (1991), and Spirtes, Glymour, and Scheines (2000 (1993)) the fact that empirical data often considerably underdetermines causal inferences – especially when it comes to inferences to complex causal structures – has become a widely recognized and investigated problem in the literature on (algorithmic) causal reasoning. All of these studies endorse a theoretical framework according to which causal structures can be analyzed in terms of Bayesian networks.[1] This framework has meanwhile become the dominant approach to algorithmic causal reasoning in the philosophical literature. Algorithms designed to discover causal Bayesian networks – henceforth *BN-algorithms*, for short – analyze probabilistic input data, i.e. probability distributions that are, for instance, acquired from frequency distributions. As section 2 is going to illustrate, the mapping of causal structures to probability distributions is not generally unambiguous. In many cases more than one causal structure is assigned to one probability distribution by BN-algorithms. Such ambiguities are not normally considered to be particularly surprising or worrisome in the literature, for, clearly, the causal inferences licensed by corresponding empirical data crucially hinges on the latter's quality, which in case of probabilistic data, as is well known, can be negatively affected by ever so many factors. For instance, frequency distributions may feature a considerable amount of confounding noise. As long as not all relevant factors involved in an investigated causal structure are controlled for in the set-up of a pertaining study, corresponding data tends to be confounded by hidden variables and

---

[1] As regards the notion of a Bayesian network cf. e.g. Pearl (1985).

is, hence, likely not to unambiguously reflect underlying causal structures. If ambiguities in causal reasoning can at least partially be ascribed to confounding noise in the data, the ambiguity ratio of a particular frequency distribution can simply be understood as an indicator of how close the set-up behind a corresponding study has come to an ideal noise-free set-up.

This raises the question as to whether it is possible to unambiguously infer complex causal structures from data that has been collected against (idealized) homogeneous causal backgrounds without confounding noise, i.e. from data that directly exhibits *deterministic* dependencies among specific configurations of variables, or whether inferences to causal structures – both of the probabilistic or the deterministic type – are underdetermined by empirical data in principle. This is one of the core questions addressed in this paper. While, since the early 1990ies, all the studies concerned with ambiguous causal inferences have been focussing on ambiguities affecting the causal interpretation of probabilistic data, this paper investigates whether similar ambiguities arise in case of deterministic systems. Addressing this question is not only relevant for determining whether the empirical underdetermination of causal inferences is a matter of principle or merely a consequence of unsuitable or noisy data. It is also of relevance to those areas of causal research that are in fact analyzing deterministic data, such as studies conducted in fully controlled laboratory contexts or so-called *narrow case* or *small-N studies* in social sciences.[2]

As is well known, algorithms uncovering causal Bayesian networks are not applicable to deterministic structures that are investigated on a sufficiently fine-grained level such that deterministic dependencies actually show up in the data, because such systems violate one of the fundamental assumptions of the BN-framework: causal faithfulness.[3] Before we can thus address the primary question of this paper, the basics of a Boolean procedure of causal reasoning, which is custom-built for the analysis of deterministic structures, are introduced in section 3. Section 4 then shows that even data that has been collected against a noise-free causal background such that deterministic dependencies are exhibited cannot always unambiguously be assigned to causal structures. In fact, we shall find that there exists a systematic empirical underdetermination of inferences to one specific causal structure: deterministic chains. It will turn out that whenever deterministic data could be modeled in terms of a causal chain, it could equally be modeled in terms of a common cause structure. To every deterministic chain there exists an empirically equivalent common cause structure. This is what I shall dub the *causal chain problem*. The paper concludes by indicating an important asymmetry between modeling deterministic data in terms of a chain and a common cause structure, respectively. Even though the deterministic chain and common cause

---

[2] Cf. Ragin (1987), Ragin (2000), or Mahoney (2000).

[3] Cf. e.g. Spirtes, Glymour, and Scheines (2000 (1993)), pp. 53-57, Glymour (2007). The causal faithfulness assumption is also briefly reviewed in section 2 below.

models are empirically equivalent they behave radically differently under model expansions.

2 Probabilistic Indistinguishability

In order to determine how causal structures and probability distributions are connected, all the different causal discovery algorithms developed in the BN-framework impose two important constraints on the causal structures and the probabilistic data generated by these structures: The structures and the data must satisfy the conditions expressed in the *causal Markov assumption* and in the *faithfulness assumption*.[4] The causal Markov assumption states that in a probability distribution $\mathcal{P}$ generated by a (acyclic)[5] causal structure $\mathcal{S}$ a variable Z is independent of all its non-effects in $\mathcal{S}$ conditional on all of Z's direct causes, provided that no direct common causes of any two variables in $\mathcal{S}$ are left out of $\mathcal{P}$. According to the faithfulness assumption, there are no other conditional independence relations in $\mathcal{P}$ than the ones implied by the causal Markov assumption. As this section is going to briefly review, probability distributions are not unambiguously connected to causal structures by these two assumptions.

Such as to illustrate the ambiguities that arise when it comes to causally analyzing probabilistic data consider the structures (a), (b) and (c) graphed in figure 1. If these structures are causally Markov and faithful, they generate
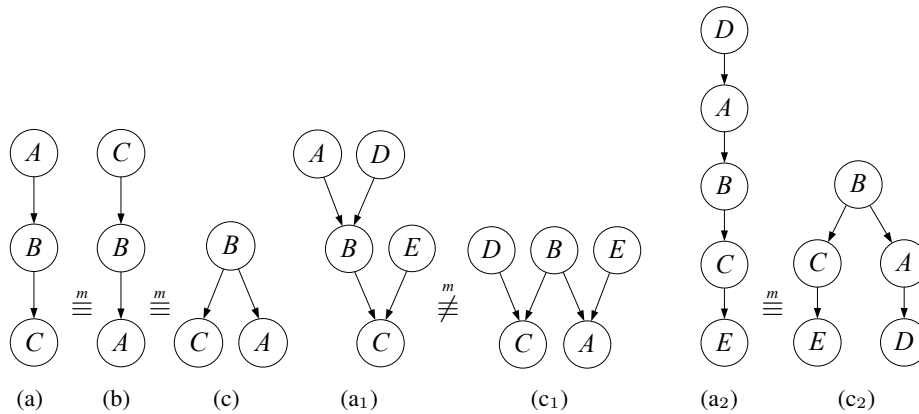

Fig. 1: (a), (b), and (c) represent three Markov equivalent causal structures – that equivalence being symbolized by "$\overset{m}{\equiv}$". Suitable extensions of graphs (a) and (c) yield graphs (a$_1$) and (c$_1$), which are not Markov equivalent. Not all extensions of (a), (b) and (c), however, resolve equivalencies. This is illustrated by extensions (a$_2$) and (c$_2$).

---

[4] Cf. e.g. Spirtes, Glymour, and Scheines (2000 (1993)), pp. 29-31, Glymour (1997), Glymour (2007).

[5] Both Boolean and BN-methodologies of causal reasoning are designed to uncover acyclic structures only. The causal structures considered in the following are hence implicitly assumed not to feature feedbacks.

probability distributions that exhibit identical conditional independence relations. The three variables or factors contained in these structures are positively correlated except for A and C being independent given B in all three structures, i.e. $p(A \mid B \wedge C) = p(A \mid B)$. (a), (b) and (c), hence, are *Markov equivalent* or *probabilistically indistinguishable*, respectively.[6] They constitute a *Markov equivalence class*. Structures in a Markov equivalence class cannot be discerned by BN-algorithms.

If indistinguishable structures are suitably extended by further variables, ambiguities may disappear. For instance, structures (a_1) and (c_1) in figure 1 are such disambiguating extensions of (a) and (c), respectively. Provided that the structure depicted in graph (c_1) is causally Markov and faithful, it, for example, gives rise to a dependence of A and E, i.e. $p(A \mid E) > p(A)$, which is not the case for the structure represented by (a_1). However, not all extensions of Markov equivalents resolve ambiguities, as e.g. extensions (a_2) and (c_2) of (a) and (c) illustrate. Even though (a_2) and (c_2) are probabilistically distinguishable from (a) and (c), they are again contained in one and the same Markov equivalence class, as they induce identical conditional independence relations among their five variables. Frydenberg (1990) and Verma and Pearl (1991) have shown that probabilistic indistinguishability can be nicely captured in graphical terms: Two (acyclic) causal graphs G_1 and G_2 represent two probabilistically indistinguishable causally Markov and faithful causal structures iff (i) G_1 and G_2 have the same vertex set; (ii) G_1 and G_2 have the same adjacencies; and (iii) G_1 and G_2 have the same *unshielded colliders*, where an unshielded collider is constituted by two edges that collide at the same vertex and whose tails are not adjacent.[7]

Countless causal structures are probabilistically indistinguishable in this sense and, accordingly, many causal inferences drawn on the basis of probabilistic data are ambiguous. There are several proposal in the literature as to how to disambiguate the causal interpretation of probabilistic data. Broadly, these proposals can be grouped into two categories. The first category is constituted by suggestions to the effect that structures as (a), (b) and (c) or (a_2) and (c_2) should be discerned by drawing on additional empirical information as temporal orderings of token events that instantiate the event types involved in these structures.[8] The disambiguation approaches contained in the second category resort to established causal knowledge about the specific structure under investigation, as e.g. knowledge about how to manipulate a particular structure in a systematic way.[9]

---

[6] Cf. e.g. Pearl (2000), pp. 19, 145, Verma and Pearl (1991).

[7] Cf. also Spirtes, Glymour, and Scheines (2000 (1993)) and Glymour (1997). Two vertices V_1 and V_2 are said to be *adjacent* in a graph G iff there is an edge between V_1 and V_2 in G. In a directed edge from a vertex V_1 to a vertex V_2, V_1 is called the *tail* and V_2 the *head*. Two edges collide at a vertex V_1 iff V_1 is the head of both edges.

[8] Cf. e.g. Suppes (1970).

[9] Cf. Spirtes, Glymour, and Scheines (2000 (1993)), ch. 4, Pearl (2000), or Woodward (2003), ch. 3.

As indicated in the introduction, rather than further investigating how causal inferences based on probability distributions can be disambiguated, the paper at hand shall be concerned with the question whether similar ambiguities arise in case of deterministic structures. To this end, the next section introduces the basics of a Boolean procedure that uncovers such structures.

3 Uncovering Deterministic Structures

Even though there is little disagreement in the literature on the fact that many causal structures, especially on a macro level, are ultimately deterministic, deterministic data cannot be analyzed within the dominant (BN) framework, as such data features more conditional independence relations than are implied by the Markov condition and, thus, violates faithfulness.[10] In order to illustrate this violation of faithfulness consider structure (c) in figure 1 and, for simplicity, suppose that B is sufficient and necessary for C. In that case, B and A are independent conditional on C, i.e. $p(A|B \wedge C) = p(A|C)$, which is not implied by the causal Markov assumption. In a deterministic structure every value of at least one exogenous variable uniquely determines the values of at least one endogenous variable.[11] Such deterministic dependencies may, of course, not show up in corresponding data, if, for instance, not all variables involved in the structure are contained in the set of investigated variables or if not all relevant factors are controlled for in a pertaining study. However, if deterministic structures are investigated against a causally homogeneous background – say, in a laboratory context – to the effect that deterministic dependencies are actually exhibited in the data, the faithfulness assumption is violated as illustrated above.

Deterministic causal structures are traditionally analyzed by methodologies invoking Boolean techniques as e.g. Quine-McCluskey optimization of truth-functions.[12] The main modern developments in this tradition can be found in Ragin (1987, 2000) and May (1999), as well as in Baumgartner (2008). For lack of space, a Boolean algorithm cannot be fully exhibited in the present context. This section is merely going to present the essential conceptual core of Boolean causal reasoning. The exemplary data used to illustrate the causal chain problem will then be kept as simple as possible, such that the problem becomes transparent without in-depth insights into the details of Boolean algorithms.

Boolean methodologies are designed to unfold deterministic structures on type level, i.e. they analyze general causation.[13] The relata of general causation are event types or *factors* for short. A factor that causes another factor is said to be

---

[10] Cf. e.g. Spirtes, Glymour, and Scheines (2000 (1993)), pp. 53-57, or Glymour (2007).

[11] For details cf. Glymour (2007), p. 236.

[12] Cf. e.g. in Quine (1952, 1959).

[13] At least two kinds of causal relations must be discerned: "Drinking is a cause of drunkenness" is a case of *general causation*, i.e. causation among event types or factors, while "Shamus' drinking of 6 beers at noon on September 7, 2004 causes Shamus' drunkenness in the afternoon of September 7, 2004" relates token events and, accordingly, is a case of *singular causation*.

causally relevant to the latter. Factors are taken to be similarity sets of event tokens. They are sets of type identical token events, of events that share at least one feature. Whenever a member of a similarity set that corresponds to an event type occurs, the latter is said to be *instantiated*. Factors are symbolized by italicized capital letters A, B, etc. 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. That means Boolean algorithms are custom-built for deterministic structures featuring binary variables.[14]

Causal analyses are always relativized to a set of investigated factors. This set is referred to as the *factor frame* of a causal investigation. 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 close-by or *coincidently*. Moreover, causes do not determine their direct effects to occur anywhere and anytime, but close-by. Determining a specific spatiotemporal interval such that, when factors are instantiated within that interval, they can be said to be coincidently instantiated is an intricate problem that, for lack of space, shall be sidestepped here.[15] As is usually done in studies on causal reasoning, I shall simply assume that for a given causal process under investigation it is sufficiently clear what the coincidence relation amounts to. Coincidently instantiated factors are termed *coincidences*. A coincidence can be seen as 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 pertaining factors: $A_1 A_2 \ldots A_n$. Coincidences constitute the empirical data processed by Boolean procedures.[16] Data collection, of course, has to comply with specific constraints that guarantee the causal interpretability of that data. For instance, data must be collected against a homogeneous background or data collection must be exhaustive. As we are exclusively concerned with the causal interpretation of empirical data, the latter can simply be assumed to be properly collected here.

Such as to illustrate the nature of the data processed by Boolean procedures, take the factor frame $\mathcal{F}_1$ consisting of the factors A, B, C, D, and E. Suppose, the behavior of these five factors is regulated by some deterministic causal structure. That there are deterministic dependencies among these factors means that they are not co-instantiatable in all logically possible combinations. If, for instance, factor

---

[14] The restriction to binary variables primarily serves conceptual simplicity. It allows for a straightforward implementation of Boolean optimization procedures, which are of great relevance to the uncovering of deterministic structures. Nonetheless, the restriction to binary variables implies that structures involving multi-valued variables must be encoded in binary terms before they can be treated by Boolean procedures, which, though always possible in case of deterministic structures, comes at the cost of an increased computational complexity (cf. e.g. Brayton and Khatri (1999)).

[15] For more details on the notion of coincidence cf. Baumgartner ((2008)), appendix A.

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

$$ABCDE$$
$$ABC\overline{D}E$$
$$A\overline{B}CDE$$
$$A\overline{B}C\overline{D}E$$
$$\overline{A}BCDE$$
$$\overline{A}BC\overline{D}E$$
$$\overline{A}\,\overline{B}C\,DE$$
$$\overline{A}\,\overline{B}\,\overline{C}\,\overline{D}\,\overline{E}$$

Tab. 1: Exemplary coincidence list to be analyzed by Boolean methodologies.

A determines factor C, it is empirically impossible to realize the coincidence $A\overline{C}$. Thus, in order to infer a deterministic structure from empirical data, we need to know what coincidences of the factors in our investigated frame are empirically realizable. Boolean procedures, accordingly, infer causal structures from lists of e.g. experimentally observed coincidences as listed in table 1. The first row of that exemplary *coincidence list* features the coincidence $ABCDE$, which means that the five factors in $\mathcal{F}_1$ have been observed to be instantiated coincidently. The second row then reports an observation of $ABC\overline{D}E$ etc. In a nutshell, Boolean algorithms assign causal structures to such coincidence lists by determining sufficiency and necessity relationships among the factors contained in the lists and by suitably minimalizing these dependencies in order to eliminate redundancies.[17]

For example, relative to the data listed in table 1 the coincidence $ABCD$ is sufficient for E, because this list contains one row such that $ABCD$ is instantiated in combination with E, *viz.* the first row, and no row such that $ABCD$ is given and E is not. $ABCD$, however, is not *minimally sufficient* for E, for $ABCD$ contains proper parts that are themselves sufficient for E. A conjunction of coincidently instantiated factors $A_1 A_2 \ldots A_n$, $n \leq 1$, is a minimally sufficient condition of a factor B iff $A_1 A_2 \ldots A_n$ is sufficient for B, i.e. $A_1 A_2 \ldots A_n \to B$, and there is no proper part α of $A_1 A_2 \ldots A_n$ such that $\alpha \to B$. A *proper part* of a conjunction designates the result of any reduction of this conjunction by one conjunct. If this notion of a minimally sufficient condition is applied to table 1, we find, for instance, that ABC – which is a proper part of $ABCD$ – is also sufficient for E, for table 1 does not record a coincidence such that ABC is given and E is not. ABC still contains sufficient proper parts: All three of its conjuncts are themselves sufficient for E relative to table 1. If we let a Boolean discovery algorithm parse through table 1 in this vein and identify minimalized deterministic dependencies, we get the following overall result:[18]

$$A \lor B \leftrightarrow C$$
$$A \lor B \lor C \lor D \leftrightarrow E \tag{R}$$

---

[17] For further details cf. Baumgartner (unpublished) and Baumgartner (2008).

[18] Of course, C and E are moreover each minimally sufficient for themselves. However, as self-causation is normally excluded, these reflexive dependencies are not amenable to a causal interpretation to begin with. Reflexive dependencies are therefore neglected in the context at hand.
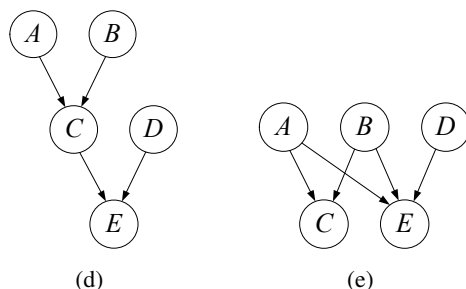
Fig. 2: A causal chain and an epiphenomenon that both could underly the behavior of the factors in $\mathcal{F}_1$ and, accordingly, generate the data in table 1.[19]

In our exemplary coincidence list factor C is instantiated if and only if A or B are instantiated, and E is instantiated if and only if A, B, C or D are instantiated. This follows from the fact that there are no coincidences including $A\overline{C}$, $B\overline{C}$, $A\overline{E}$, $B\overline{E}$, $C\overline{E}$, $D\overline{E}$, $\overline{ABC}$ and $\overline{ABCDE}$ in table 1. That does not mean that, say, A is sufficient for C in isolation. As mentioned above, causally interpretable data must always be collected against a particular background. Correspondingly, table 1 merely determines A to be sufficient for C relative to that specific configuration of unanalyzed background conditions.

A Boolean procedure first identifies a set of deterministic dependencies – (R) in our simple exemplary case – that hold among the factors in an analyzed coincidence list and then, in a second step, causally interprets that set. Without going into the details of causal interpretation rules, it can easily be seen that the data in table 1 could have been generated by, at least, the two causal structures depicted in figure 2.[19] Both the deterministic chain (d) and the common cause structure or *epiphenomenon*[20] (e) give rise to exactly those dependencies among the factors in $\mathcal{F}_1$ that are reported in (R). Or put differently, according to both (d) and (e) precisely those coincidences of A, B, C, D, and E are empirically realizable which

---

[19] Dependencies as recorded in (R) are not as straightforwardly causally interpretable as might be suggested here. Mackie's (1974) famous Manchester Factory Hooters example demonstrates that minimally sufficient conditions are not directly amenable to a causal interpretation. In Baumgartner (2008) I indicate what additional constraints have to be met in order to warrantably causally interpret dependencies as in (R). For the context at hand, however, we can ignore these complications.

[19] Note that the arrows in these graphs represent a *specific* form of deterministic direct causal relevance: An edge as $\langle A, C \rangle$ signifies that A is a sufficient direct cause of C. This graphical notation is not to be confounded with arrows in graphical representations as can be found in Spirtes (2000 (1993)) or Pearl (2000) where arrows stand for *any* functional dependence.

[20] In the following, I interchangeably speak of *common cause structures* and *epiphenomena*. Note that this terminology differs from the notion of an epiphenomenon used in the literature on mental causation. In the latter context an epiphenomenon is a physically caused mental side effect which itself cannot cause anything. Here "epiphenomenon" just describes a causal structure featuring at least one cause with at least two parallel effects (cf. e.g. graphs (c), (c$_1$) or (c$_2$) of figure 1 or (e) of figure 2). Nothing with respect to a causal impotence of these parallel effects is implied by referring to such a structure as being *epiphenomenal*.

are listed in table 1. If these two deterministic structures are investigated against a causally homogeneous background, they generate the exact same coincidence data. I shall, therefore, refer to (d) and (e) as *c-equivalent* causal structures. In sum, thus, Boolean algorithms assign sets of *c*-equivalent causal structures to coincidence lists as table 1 along the lines sketched in this section. These sets sometimes contain one single structure, and sometimes, as illustrated in the example discussed in this section, multiple *c*-equivalent structures.[21]

4   The Problem

Notwithstanding the fact that structures (d) and (e) regulate the behavior of the factors in $\mathcal{F}_1$ identically, they differ in causal respects. According to (d), A is *indirectly* causally relevant for E, while it is attributed *direct* causal relevance for E by (e). Moreover, C is causally relevant for E in (d), yet not in (e). These significant differences in causal structuring, however, do not affect the behavior of the factors in $\mathcal{F}_1$ at all. Any coincidence of these factors is either compatible with both (d) and (e) or with neither of the two structures. This shows that not only probabilistic data, but also (noise-free) deterministic data is not always unambiguously causally interpretable. Furthermore, this section is going to demonstrate that the *c*-equivalence of (d) and (e) is not some idiosyncrasy of these particular structures or of the data listed in table 1. Rather, to *every* deterministic chain there exists an epiphenomenon generating the same coincidence data as the chain. Every deterministic chain is reducible to an epiphenomenon.

In order to see the general reducibility of chains to epiphenomena, it suffices to realize that the characteristic structural feature of deterministic chains can also be found in a specific kind of epiphenomena. Every deterministic chain comprises at least two *entangled factors*: Two factors X_1 and X_2 are entangled iff all factors contained in minimally sufficient conditions of X_1 are part of minimally sufficient conditions of X_2 as well. A factor X is said to be *part* of a minimally sufficient condition α if X is a conjunct contained in α. If X is the only conjunct in a minimally sufficient condition, it is trivially part thereof. To parts of minimally sufficient conditions I shall also refer as *determinants* in the following. While every deterministic chain features at least two factors X_1 and X_2 such that all determinants of X_1 are determinants of X_2 and, as a logical consequence thereof, X_1 is a determinant of X_2, some common cause structures also comprise thus entangled factors. To illustrate, consider our exemplary structures (d) and (e): Factors C and E are not only entangled in (d) but also in (e). Whatever is minimally sufficient for C is also minimally sufficient for E in both (d) and (e). Epiphenomena featuring at least two entangled factors shall, accordingly, be referred to as *entangled epiphenomena*. The chain (d) can be claimed to exhibit the dependencies expressed in (1), while the epiphenomenon (e) specifies the dependencies in (2).

---

[21] Cf. Baumgartner (2008).

$$(A \lor B \leftrightarrow C) \land (C \lor D \leftrightarrow E) \qquad (1)$$

$$(A \lor B \leftrightarrow C) \land (A \lor B \lor D \leftrightarrow E) \qquad (2)$$

(1) and (2) are logically equivalent, which can easily be established by alternately substituting C and $A \lor B$ in the second conjuncts of the two expressions. This substitutability is guaranteed by the first conjuncts of (1) and (2). The two expressions are moreover logically equivalent to a conjunction of the dependencies mentioned in (R). To every expression of type (1) there exists a logically equivalent expression of type (2), and, accordingly, to every deterministic chain there exists a c-equivalent entangled epiphenomenon.[22]

While entanglements are the characteristic of chains, there are epiphenomena without entangled factors, as e.g. structure (c_1) in figure 1. Epiphenomena without entanglements are not c-equivalent to a chain. According to (c_1), the parallel effects A and C of the common cause B can be instantiated in all logically possible combinations, whereas in virtue of the corresponding chain (a_1) C is instantiated whenever there is an instance of A. That is, while all chains contain at least two entangled factors, epiphenomena may, but must not feature entanglements. 'Ordinary' deterministic epiphenomena without entanglements are unambiguously identifiable via the coincidence lists they generate.

However, based on deterministic coincidence data an unambiguous inference to causal chains is excluded in principle. Coincidence data which can be modeled in terms of a chain can equally be modeled in terms of an entangled epiphenomenon. To this indistinguishability of deterministic chains and entangled epiphenomena I shall in the following refer as the *chain problem*.

Chain Problem: Whenever causal structures are investigated on such a fine-grained level that deterministic dependencies are exhibited in pertaining coincidence data $\mathcal{L}$ and $\mathcal{L}$ is compatible with a causal chain $\mathcal{S}_1$, there exists an entangled epiphenomenon $\mathcal{S}_2$ such that $\mathcal{L}$ is compatible with $\mathcal{S}_2$ as well. Deterministic (noise-free) coincidence data cannot unambiguously be identified to be the result of a chain.

Note that the chain problem is characteristic for deterministic data, i.e. data that does not exhibit all 2^n logically possible configurations of n factors in a causally analyzed frame. If the causal structure behind our exemplary coincidence list 1 were not investigated against a noise-free background to the effect that our data would consist of a frequency distribution listing frequencies for all 32 logically possible combinations of the five factors A, B, C, D, E and if that frequency distribution, moreover, were to satisfy the Markov and faithfulness assumptions, BN-algorithms would be able to distinguish between structures (d) and (e). As (d) and (e) do not have identical unshielded colliders they are not Markov equivalent.

---

[22] Further exemplary reductions of chains to epiphenomena can be found in figures 4, 5, and 6 below.

This finding, of course, raises the somewhat counterintuitive question whether deterministic structures should deliberately not be investigated on such a fine-grained level that deterministic dependencies are manifested, i.e. whether a certain amount of confounding noise is a desirable feature of causally analyzed empirical data after all. For, while noise-free deterministic data in principle does not allow for unambiguous inferences to chains, unambiguous inferences to chains are not excluded relative to a noisy background that generates a causally Markov and faithful frequency distribution. An answer to this question has to await another paper.

The rest of this paper shall be dedicated to the question how ambiguities affecting the causal interpretation of deterministic data can possibly be reduced or even avoided completely. For, clearly, the chain problem heavily conflicts with common intuitions according to which deterministic chains are omnipresent in nature, at least on a macro level. Hence, what are the grounds based on which we ever so often intuitively or pre-theoretically model deterministic data in terms of chains rather than common cause structures? This question becomes even more pressing if we no longer treat the factors in our exemplary frame $\mathcal{F}_1$ in terms of abstract variables, but provide them with concrete interpretations. For as soon as such interpretations are provided, causal intuitions are as firm as can be when it comes to opting for the chain or the epiphenomenon given a coincidence list as in table 1. Consider the following interpretations:

Interpretation (I): Assume a car engine can be started in two ways only: either by turning the key in the starter lock or by short-circuiting the ignition cable. Whenever the engine is running, the corresponding car begins to move. The car can be set in motion by alternative factors also, such as towing or pushing, i.e. by external impulses:

A = Turning the key in the starter lock
B = Short-circuiting the ignition cable
C = Running engine
D = External impulse
E = Motion of the wheels.

Interpretation (I) clearly suggests the underlying causal structure to be (d) in figure 2. Hence, provided that the factors in table 1 are interpreted according to (I), we tend to model the underlying process in terms of a chain.

Interpretation (II): Suppose in a particular city there are exactly two power stations. The power supply of a specific house, say house a, in that city entirely depends on the power production in at least one of the two stations. Another house, call it b, is equipped with a generator for cases of citywide power failures. Whenever one of the two power stations produces electricity, both a and b are power supplied:

$$
\begin{aligned}
A &= \text{Power production by station 1} \\
B &= \text{Power production by station 2} \\
C &= \text{Power supply of house } a \\
D &= \text{Power production by the generator in } b \\
E &= \text{Power supply of house } b.
\end{aligned}
$$

If the coincidences in table 1 are interpreted in the vein of (II), the behavior of the factors in $\mathcal{F}_1$ is intuitively seen to be regulated by a common cause structure of type (e).

Intuitively there is no doubt that the causal process starting with turning the key in the starter lock and resulting in the motion of the wheels has the form of a chain and that there is a common cause structure behind the power supply of the houses a and b. Still, the coincidences in table 1 alone neither warrant the first nor the second of these inferences. Therefore, our firm intuitions as regards causal modeling relative to a respective interpretation cannot exclusively be based on the deterministic data contained in table 1.

5   Disambiguation Candidates

Prima facie, a plausible reaction to the chain problem as presented thus far will be to claim that this problem merely demonstrates the impossibility to positively identify deterministic chains based on coincidence information alone. The chain problem might be seen to establish that coincidence data must be complemented by additional empirical information in order to allow for unambiguous inferences to chainlike structures. When we pre-theoretically or informally model the processes behind interpretations (I) and (II) we seem to implicitly draw on additional information, which apparently resolves c-equivalencies. There are several conceivable sources of additional information based on which solutions of the chain problem might be developed.[23] Let us consider them in turn.

5.1   Chronological Order

One suggestion to solve the chain problem by consulting additional sources of empirical information could be to impose a chronological ordering onto the instances of the factors in a causal structure. If it is stipulated that causes always occur *before* their effects, a chain structure as graph (d) in figure 2 could be claimed to determine instances of C to occur *prior to* the instances of E. This, in turn, does not hold for the epiphenomenal structure depicted in graph (e). An entangled epiphenomenon as (e) is compatible with instances of C and E occurring *simultaneously*. Such as to illustrate this difference, graphs (d) and (e) are confronted with a timeline in figure 3.

---

[23] Most of the suggestions as to how to disambiguate inferences to complex causal structures considered in the following have also been discussed in the context of resolving probabilistic ambiguities (cf. e.g. Suppes (1970), Spirtes, Glymour, and Scheines (2000 (1993)), Hausman (1998), or Woodward (2003)).
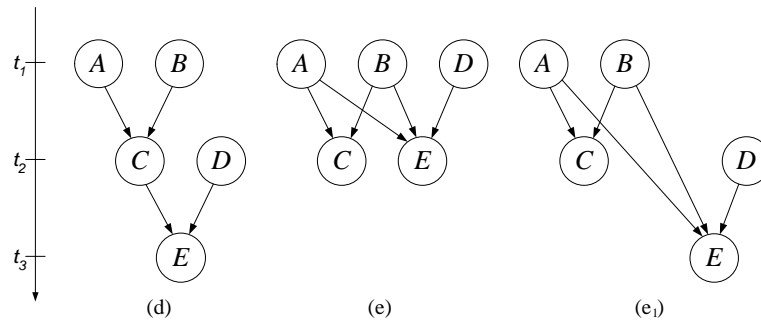
Fig. 3: The arrow on the left hand side represents the direction of time such that $t_1 < t_2 < t_3$. While (e) determines instances of C and E to occur simultaneously, (d) and (e$_1$) represent causal structures according to which C is instantiated before E.

As is well known, it is highly dubitable whether causes in fact can justifiably be claimed to always occur before their effects and, thus, whether a simultaneity of causes and effects can be excluded on a priori conceptual grounds.[24] Nevertheless, such as to see whether the ambiguities induced by the chain problem could, in principle, be solved along these lines, let us abstain from these difficulties in the present context and assume that causes generally occur prior to their effects. Subject to that assumption, factors involved in a coincidence list as the one in table 1, which could stem both from a chain and an entangled epiphenomenon, have to be compared as to the chronological ordering of their instances. Finding C and E to be simultaneously instantiated then prompts an inference to the epiphenomenal model. So far so good. Yet, what if instances of C and E are not found to occur simultaneously? Can the structure behind table 1 unambiguously be modeled in terms of a chain, if it is found that events of type C occur prior to events of type E? As structure (e$_1$) in figure 3 shows, that is not the case. Epiphenomenal structures are *compatible* with simultaneous occurrences of their parallel effects, but they do *not determine* such simultaneity. Effects of a common cause might well be instantiated sequentially. Suppose, houses *a* and *b* in the power station example are connected to the power stations by wiring of different length or of different conductivity, such that electricity always reaches house *a* prior to house *b*. Nonetheless, of course, the causal structure behind this chronological specification of the power station example is to be modeled in terms of an epiphenomenon. Hence, while (e) might be identifiable by means of chronological constraints, (d) and (e$_1$) cannot thus be kept apart, for they are not only *c*-equivalent, but also chronologically equivalent. Every coincidence list that could be the result of a chain might just as well be the product of a chronologically ordered entangled epiphenomenon of type (e$_1$). Building the direction of time into a criterion that distinguishes between chains and entangled epiphenomena would merely allow for identifying those entangled epiphenomena that happen to be constituted by simultaneously occurring parallel effects. Such a

---

[24] Cf. e.g. Lewis (1979), Brand (1980) or Huemer and Kovitz (2003).

criterion would, however, be of no help when it comes to identifying deterministic chains.

5.2   Spatiotemporal Proximity

The direction of time is not the only further empirical information that could be resorted to as a means to distinguish between chains and entangled epiphenomena. Spatiotemporal proximity might be proposed as an alternative. If it is stipulated that instances of causes occur proximately to the instances of their direct effects, a chain structure as graph (d) can be claimed to determine instances of C to occur proximately to the instances of E. This does not hold for the epiphenomenal structure (e). On the contrary, in an epiphenomenon there is no necessary spatiotemporal connection between instances of parallel effects of a common cause.

As mentioned in section 3, causes determine their direct effects to occur within a certain spatiotemporal interval or within the same situation. Yet, it is far from clear what exactly this spatiotemporal relation is. Depending on the causal process under investigation and the level of specification chosen for a causal analysis, instances of causes and their direct effects can be said to be properly related only if they are in plain spatiotemporal contact, while in other cases causes may well be instantiated some nonzero distance away from their direct effects. The allowable distance between causes and effects cannot be fixed to a specific spatiotemporal interval. The theory of Special Relativity only provides an upper bound: Causes and effects must be instantiated within each other's light cones. Notwithstanding this lacking specificity, given a concrete causal process and a level of specificity chosen for the analysis of that process it is normally uncontroversial which factors can be said to be proximately instantiated. In order to see whether the chain problem could, at least in principle, be solved by somehow drawing on the spatiotemporal association of the instances of causes and effects, let us, thus, in the present context ignore the notorious fuzziness involved in the notion of spatiotemporal proximity of the causal relata.

Assume we have identified a nonzero interval s that constitutes a suitable interpretation of proximity for the factors in our exemplary frame $\mathcal{F}_1$. Such as to determine whether the structure underlying the behavior of the factors in $\mathcal{F}_1$ is a chain or an epiphenomenon by drawing on s it must be checked whether (i) C and E are instantiated within s, whereas $A \vee B$ and E are not, or whether (ii) $A \vee B$ and E are instantiated within s, whereas C and E are not. Let us suppose that the scenario given in interpretation (I) is a case of type (i) and that the scenario described by interpretation (II) is of type (ii). That is, we assume that the turning of the starter key is proximate to the firing spark plug and the running engine, yet not proximate to the turning wheels, which, instead, are proximate to the running engine. Furthermore, the two power stations shall be taken to be proximate to the power supplied houses, which themselves are not proximate to each other. Relative to such a constellation the chain and the epiphenomenal model could in fact

be discerned by drawing on the spatiotemporal relation among the instances of the involved factors.

Spatiotemporal proximity, however, only paves the way towards a solution of the chain problem if we were in fact ready to model the structures behind examples (I) and (II) differently given that spatiotemporal relations changed. Hence, let us assume that some outlandish engineer designed a car whose starter lock is located on the hubcap of the right rear wheel, to the effect that the turning of this car's starter key does not happen within s of the running engine, but, rather, within s of the rolling wheels. Yet apart from this peculiarity, this unusual car resembles ordinary cars in all detail. Or more abstractly put: Suppose this car constitutes an instance of a constellation as described in (ii). Is the turning of the starter key now directly causally relevant to the motion of the wheels? Would we be ready to model the causal structure behind this car's movement in terms of an entangled epiphenomenon? Certainly not. The situation is completely analogous in case of entangled epiphenomena. Supposing that houses a and b of the power supply example are neighboring or, more generally, instantiated in terms of (i) does not change the causal structuring of the underlying process at all. Given that the two houses are proximate does not turn the epiphenomenon into a chain.

If the distinction between chains and entangled epiphenomena really hinges on some form of spatiotemporal proximity, it must be a very special kind of proximity that accounts for it. It must be a form of proximity such that 'causal influence' is transmitted from an instance of a cause to its proximate successor. In case of the car with the starter lock on its hubcap, for instance, such transfer of 'causal influence' can be claimed to proceed through the wiring and by mediation of the engine and not directly from the lock to the wheels. Yet, a notion of spatiotemporal proximity that involves transmission of 'causal influence', obviously, is of no use when it comes to inferring causal structures from empirical data without prior causal knowledge about the processes underlying the data. For whether 'causal influence' is transmitted from C to E in our exemplary frame is just what is under investigation.

5.3 Transference

Nonetheless, this finding might be taken to indicate that the chain problem could be solved in the vein of the so-called *transference theory* of causation.[25] The following criterion to distinguish between the chain and the epiphenomenal model with respect to the frame $\mathcal{F}_1$ might be proposed: If it is found that energy or momentum or, more generally, some conserved quantity is transferred from instances of factor C to instances of E, the behavior of the factors in $\mathcal{F}_1$ is regulated by a causal chain – otherwise by an entangled epiphenomenon.

Prima facie, this criterion in fact seems to perfectly capture our pre-theoretic reasons to model scenario (I) in terms of a chain and (II) in terms of an epiphe-

---

[25] For a modern variant of this account see e.g. Dowe (2000).

nomenon. There is a transfer of energy from the operating engine to the rolling
wheels, which, therefore, is a chain. In contrast, no energy is transferred from
house *a* to house *b*, which induces an epiphenomenal modeling. This modeling of
(I) and (II) is unaffected by any of the modifications of the examples discussed in
the previous section. Hence, even the causal structuring of the car whose starter
lock is located on a hubcap is identified to be a chain, because energy is transferred
from the turning of the lock to the spark plug and from the operating engine to the
wheels.

As in case of the disambiguation candidates discussed above, there are prin-
cipled objections against drawing on transfer processes in causal reasoning. First,
it has been pointed out that by far not all causal processes – notably in the field
of social sciences – in fact involve energy transfer, and second, it has been argued
that the identification of transfer processes presupposes a considerable amount of
prior causal knowledge about investigated structures.[26] At the core of transfer-
ence theories lies the distinction between genuine causal processes and so-called
pseudo-processes: While a genuine causal process features energy transmission,
a pseudo-process does not. Yet, in order to make sense of the notion of energy
transmission, clarity on what a causal process is needs to be provided. Identifying
the factors in an investigated structure whose instances are connected by energy
transfers is not independent of identifying the factors that are causally dependent.
Nonetheless, clarity on causal dependencies is just what the transference theoretic
framework intends to supply.[27] However, for the sake of the argument, let us again
ignore these conceptual hurdles in order to see whether, given that clarity on the no-
tion of a transfer process could somehow be presupposed, the chain problem could
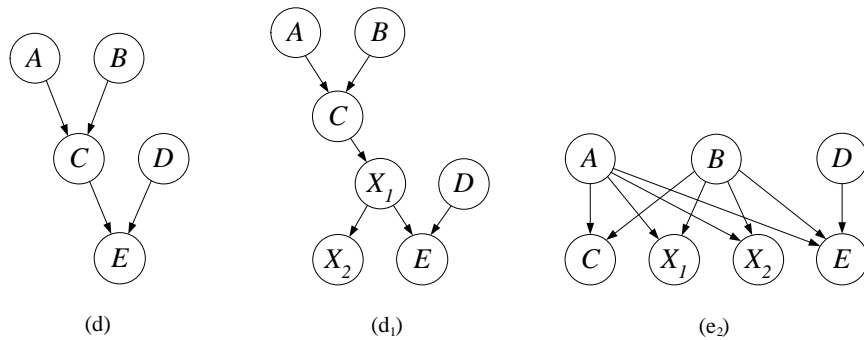at least be solved for physical processes that in fact involve such transferences.

(d)   (d_1)   (e_2)

Fig. 4: Structure (d_1) is the result of introducing two additional factors into (d): X_1 repre-
        sents a transfer process between C and E and X_2 stands for needle deflection in a
        measuring device. To (d_1) there again exists an c-equivalent epiphenomenon: (e_2).

---

[26] For a condensed presentation of the pros and cons of a transference theory of causation cf.
Dowe (Fall 2007). For more details cf. Dowe (2000), Kistler (2001).

[27] Cf. Salmon (1994).

Determining whether the structure behind the coincidences in table 1 is a chain or an epiphenomenon, according to a transfer criterion, amounts to checking if there is energy or some other conserved quantity being transferred from the instances of C to the instances of E. That, in turn, requires that some kind of measurement device is installed on the path from C to E. If this device detects a transfer process, the structure generating the data in table 1 is a chain, otherwise it turns out to be an epiphenomenon. Installing such a device induces an expansion of the causal model depicted in graph (d). In graph (d_1) of figure 4 this expansion is illustrated by the introduction of factors X_1 and X_2, where X_1 represents some conserved quantity being transferred and X_2 stands for needle deflection in a corresponding measurement device.[28] Clearly, (d_1) is not c-equivalent to the entangled epiphenomenon (e) of figure 2. Causal structures comprising different factors, trivially, do not generate identical coincidence data. Yet, as (d_1) is a causal chain as well, there again exists a c-equivalent epiphenomenon, *viz.* (e_2). Thus, while a transfer criterion may distinguish between (d_1) and (e), it does not distinguish between (d_1) and (e_2). That means claiming that the needle deflection in a measuring device is triggered by C and an ensuing transfer process, on the one hand, or by A or B, on the other, empirically amounts to the same. Now, of course further measurement devices might be installed on the edges $\langle C, X_1 \rangle$ and $\langle X_1, E \rangle$ in (d_1). However, for all such expansions of (d_1) there will exist c-equivalent epiphenomena.

These new equivalencies might be taken to show that, even if energy transfer from C to E is not measurable in a way that allows for a positive identification of a chain, it should at least be possible to test whether such a transfer process – if there is one going on – can be intercepted. For instance, without a crankshaft, which translates the reciprocating linear piston motion into rotation, the kinetic energy of the car engine is not transferred to the axis. If the crankshaft is broken, the wheels do not turn even if the engine is running. Does this finding conclusively establish the structure behind interpretation (I) to be chainlike? Intercepting the transference of kinetic energy from the engine to the axis by manipulating the crankshaft amounts to nothing else but showing that the wheels of the car only start turning if the engine is running *and* the crankshaft is working properly or there is some external impulse. Thus, C is revealed not to be sufficient for E in isolation. C determines E only in combination with a functioning crankshaft X_3: $CX_3 \vee D \leftrightarrow E$. Hence, the structure (d) must be expanded by X_3 such that C and X_3 constitute a complex cause of E. This is graphically captured by an arch

---

[28] In the social science literature there is a related methodology called *process-tracing* that aims to establish the existence of a causal mechanism between two investigated variables Y_1 and Y_2 by successively filling in intermediate variables on the path $\langle Y_1, Y_2 \rangle$ (cf. e.g. Mahoney (2000) or George and Bennett (2005), ch. 10). Some authors interested in social mechanisms argue that such mechanisms are unobservable primitive entities (cf. e.g. Steinmetz (1998)). As such they could not be treated on a par with ordinary causal variables as done in the graphs of figure 4. However, unobservable mechanisms, apparently, are of no avail when it comes to distinguish between structures (d) and (e) on empirical grounds.
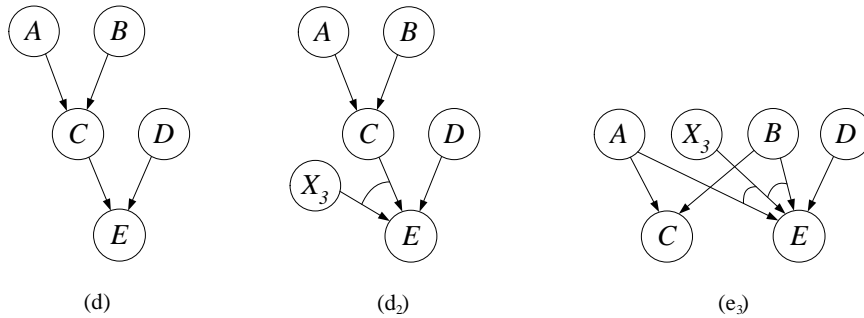
Fig. 5: (d$_2$) is the result of introducing one additional factor into (d): X_3 represents a co-factor together with which C constitutes a complex cause of E – the arch symbolizing conjunction. However, to (d$_2$) there again exists a c-equivalent epiphenomenon: (e$_3$).

connecting the edges $\langle C, E \rangle$ and $\langle X_3, E \rangle$ in graph (d$_2$) of figure 5.[29] Again, (d$_2$) and (e) clearly are not c-equivalent. However, as (d$_2$) is a chain structure as well, there exists another c-equivalent entangled epiphenomenon: (e$_3$). Whether CX_3 is said to be a deterministic complex cause of E or whether AX_3 and BX_3 are identified as E's complex causes empirically amounts to the same.

That means, even though it is possible to distinguish between two particular causal structures such as (d$_1$) and (e) or (d$_2$) and (e) by drawing on the notion of a transfer process, to every expanded chain as (d$_1$) and (d$_2$) there still exist c-equivalent epiphenomena. Thus, the chain problem is not solved by resorting to the notion of a transfer process. Rather than solving the chain problem a transfer criterion, at best, calls for further and further expansions of investigated factor frames and corresponding causal models without ever positively identifying a deterministic chain.

5.4   Interventions – Prior Causal Knowledge

As indicated in section 2, according to a popular proposal in the literature concerned with probabilistic indistinguishability, Markov equivalent structures can be discerned by suitably manipulating them. Very roughly put, e.g. Spirtes, Glymour, and Scheines (2000 (1993)), ch. 7, argue that in order to distinguish between structures (a) and (c) of figure 1 it needs to be checked whether A and C can be manipulated independently of each other.[30] If that is the case, the structure regulating the behavior of these factors is an epiphenomenon, otherwise it is a chain. In Woodward (2003), this idea is embedded into a full-fledged interventionist theory of causation. Richardson, Schulz, and Gopnik (2007) indicate that such an interventionist methodology can also be usefully implemented when it comes to uncovering

---

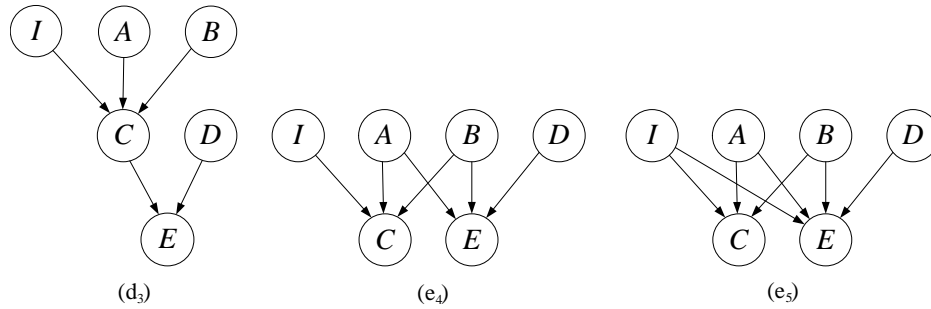[29] For details on this graphical notation cf. Baumgartner (2006), ch. 2.
[30] Similarly Pearl (2000).

Fig. 6: In structures (d_3) and (e_4) I amounts to an intervention variable for C with respect to E, whereas in (e_5) it does not.

deterministic structures. Let us therefore see whether c-equivalent structures as (d) and (e) can be kept apart by suitably drawing on interventions.

Interventionist inference procedures rest on a very specific notion of an intervention which Woodward (2003), p. 98, defines along the following lines: I is an intervention variable for a variable X with respect to a variable Y iff (i) I causes X, (ii) certain values of I render X independent of all its other causes, (iii) any directed path from I to Y goes through X, and (iv) I is (statistically) independent of any variable Z that causes Y and that is located on a directed path that does not go through X. In case of deterministic structures, an intervention variable for X with respect to Y can simply be understood as an exogenous factor I that is sufficient and directly causally relevant for X and any directed path from I to Y goes through X. For instance, I is an intervention variable for C with respect to E in structures (d_3) and (e_4) of figure 6. The variable I in structure (e_5), however, violates condition (iii) as it is not only directly relevant to C but also to E.

Drawing on that notion of an intervention in order to distinguish between structures (d) and (e) then amounts to the following: If, against the causal background of table 1, there is at least one way to intervene on C that is systematically accompanied by a change in E – for brevity, call this scenario $\mathcal{T}_1$ –, the underlying structure is a chain, otherwise – scenario $\mathcal{T}_2$ – it is an epiphenomenon.[31] Let us examine the merits of this proposal. If every manipulation of C by means of a variable I leaves E unaltered, i.e. in case of $\mathcal{T}_2$, the structure regulating the factors in our exemplary frame indeed cannot be a chain. The only structure that can accommodate an expansion of the original factor frame yielding scenario $\mathcal{T}_2$ is structure (e_4). This shows that ordinary non-entangled epiphenomena can be positively identified on empirical grounds within an interventionist framework. Does this framework also allow for a positive identification of chains? Assume that suitably manipulating C by means of a factor I is accompanied by a corresponding change in E, thus, assume that $\mathcal{T}_1$ obtains. On the face of it, there are two structures that could account for $\mathcal{T}_1$, *viz.* (d_3) and (e_5). However, only in structure (d_3) I actually satisfies

---

[31] Cf. Woodward (2003), p. 101.

the constraints imposed on interventions above. In (e_5) I violates condition (iii) because it is not only directly relevant to C but also to E. That is, if I is assumed to be an intervention variable for C with respect to E, only the chain model can accommodate a continuing dependency of C and E under manipulations of C by means of I. Provided that I is an intervention, the inference to chains and entangled epiphenomena, respectively, can in fact be disambiguated within the interventionist framework.

A positive identification of deterministic chains along these lines, of course, crucially rests on the assumption that I indeed is an intervention. How can such an assumption be justified, i.e. how is I revealed to be an intervention variable for C with respect to E given that manipulating C by means of I is accompanied by corresponding changes in E? Here a problem emerges: Demonstrating that I represents an intervention on C with respect to E relative to scenario $\mathcal{T}_1$ is not independent of determining whether the causal structure responsible for the occurrence of $\mathcal{T}_1$ is adequately modeled in terms of (d_3) or of (e_5). For in order to substantiate that I is an intervention variable for C with respect to E it must be shown that there is no direct causal connection between I and E. That means it must be shown that the structure generating scenario $\mathcal{T}_1$ is not an entangled epiphenomenon of type (e_5) – which then, of course, leaves (d_3) as only remaining model candidate. Whether that indeed is the case or not, however, is just the question raised by the chain problem. In order to establish that I is an intervention on C with respect to E, it must be presupposed that $\mathcal{T}_1$ is not adequately accounted for in terms of (e_5). Plainly, if the assumption that I represents an intervention on C with respect to E is then used to determine whether the structure regulating the behavior of the factors in the extended frame $\{A, B, C, D, E, I\}$ is of form (e_5) or not, the question raised by the chain problem is begged. In sum, a disambiguation of the inference to deterministic chains along the interventionist lines is a disambiguation by means of a causal assumption that, in case of scenario $\mathcal{T}_1$, simply excludes modeling pertaining data in terms of an entangled epiphenomenon.

It is beyond doubt, however, that ever so often in experimental practice prior causal knowledge is available that determines factors or variables to satisfy the constraints imposed on interventions above and that excludes entangled epiphenomena as possible models, respectively. In case of scenario $\mathcal{T}_1$, such prior knowledge conclusively establishes that the behavior of the factors in our exemplary frame is governed by a deterministic chain. Clearly though, along these interventionist lines the inference to chains is disambiguated on causal and not on empirical grounds. Whenever no prior causal knowledge that excludes epiphenomenal models is available, causal chains cannot be positively identified by means of interventions. Furthermore, whenever the inference to chains is disambiguated by invoking prior causal knowledge, the question arises as to what warrants that knowledge. What empirical evidence ultimately justifies assuming I to be an intervention variable for C with respect to E in case of $\mathcal{T}_1$; or, equivalently, what empirical evidence ultimately justifies the assumption that scenario $\mathcal{T}_1$ is not to be accounted for in

terms of an entangled epiphenomenon? The considerations of this section render it doubtful that, in the end, there exists an empirical vindication of such assumptions.

To sum up: All the different proposals to disambiguate the inference to deterministic chains discussed in this section embark on the same core strategy. They all aim to resolve ambiguities by broadening the empirical basis based on which the structure behind the data in table 1 is to be identified. On the one hand, we have found that even though additional information about temporal orderings or spatiotemporal relations among events instantiating the factors in table 1 may distinguish between two particular causal models as e.g. (d) and (e) in figure 3, such additional information nonetheless does not ground conclusive identifications of deterministic chains. On the other hand, resolving ambiguities by drawing on transfer processes or interventions essentially amounts to expanding the factor frame of table 1 and collecting additional coincidence information about the factors contained in those extended frames. Such collection of further coincidence data, trivially, either breaks the entanglement of C and E or it does not. The fact that C and E remain entangled relative to additional data is more precisely expressed in (CE):

(CE) For all coincidence lists $\mathcal{L}_x$ that result from expanding the factor frame $\{A, B, C, D, E\}$ of table 1 by any factors $X_1, \ldots, X_n$ and from collecting pertaining coincidence data over the frame $\{A, B, C, D, E, X_1, \ldots, X_n\}$: C and E are entangled in $\mathcal{L}_x$.

As long as collecting further coincidence data does not violate (CE), there exist at least two c-equivalent models accounting for that data equally well. Collecting more coincidence information does not by itself, i.e. without accompanying causal assumptions that exclude the rivalling epiphenomenal model, allow for an unambiguous inference to the deterministic chain. In whatever way further empirical information that complies to (CE) is brought to bear, there will always exist a common cause model that accounts for the data equally well as the chain model. The fact that interventionist methodologies perform best when it comes to unambiguously assign a causal structure to table 1 suggests that the inference to deterministic chains inevitably presupposes prior causal knowledge about the structure under investigation. The chain problem, in turn, indicates that in case of deterministic chains such knowledge is likely not to be justifiable on purely empirical grounds and, accordingly, must ultimately be grounded in (non-empirical) causal assumptions. The next section is going to reveal an important asymmetry between modeling data in terms of deterministic chains and entangled epiphenomena – an asymmetry which might be resorted to in order to justify such causal assumptions.

6 An Asymmetry Between Chains and Entangled Epiphenomena

While factor frame expansions and collecting further coincidence information about the extended frame do not ground *empirical* distinctions between deterministic chains and entangled epiphenomena, this section is going to show that such

expansions still bring about an important asymmetry between modeling data as given in table 1 in terms of a chain and an entangled epiphenomenon, respectively. In both our exemplary interpretations (I) and (II) of that data an utterly simplistic causal structure is assumed to regulate the behavior of the factors in $\mathcal{F}_1$. Real cars can be started in manifold ways, each of which involve many more factors than merely the turning of a starter key or the short-circuiting of a cable. The same holds for power supplied houses. The mere operating of power stations is commonly neither sufficient nor necessary for the power supply in houses. It takes a sophisticated infrastructure to bring the electricity to the consumers and generators or batteries can supply any houses, not just b, with power. $\mathcal{F}_1$ is extendable in manifold ways when it comes to generating a more realistic scenario with respect to our examples. Against the background of interpretation (I), $\mathcal{F}_1$ is extendable by, say, the factors:

G = Starting the engine by use of jumper cables
H = Starting the engine by use of a mechanical starting device
I = Transmission of an electrical impulse to the spark plug.

The thus extended factor frame shall be labeled $\mathcal{F}_2$. In contrast to interpretation (I), $\mathcal{F}_1$, when interpreted according to (II), is e.g. extendable by the following factor:

L = Power production by the generator in a.

This expansion of $\mathcal{F}_1$, interpreted in terms of (II), will be referred to as $\mathcal{F}_3$.

The expansions of $\mathcal{F}_1$ to $\mathcal{F}_2$ and $\mathcal{F}_3$ differ in an important respect: The expansion of $\mathcal{F}_1$ to $\mathcal{F}_2$ is *structure-conserving*, whereas the same does not hold for the expansion of $\mathcal{F}_1$ to $\mathcal{F}_3$. The feature of the graphs (d) and (e) which is characteristic for the chain problem is the entanglement of C and E. This entanglement is conserved when $\mathcal{F}_1$ – interpreted along the lines of (I) – is expanded to $\mathcal{F}_2$. It is not only conserved upon expanding $\mathcal{F}_1$ by G, H, and I. *Every* factor that is newly introduced into a deterministic chain as (d) and that is part of a minimally sufficient condition of C, is part of a minimally sufficient condition of E as well. Whatever determinant of C is introduced into structure (d) above the C vertex, it is a determinant of E as well. This follows from the mere structure of causal chains. In contrast to the structure-conserving expansion $\mathcal{F}_2$ of $\mathcal{F}_1$, expanding $\mathcal{F}_1$ – interpreted along the lines of (II) – to $\mathcal{F}_3$ *suspends the entanglement* of C and E. For relative to such an expansion there will be cases in which house a is power supplied while house b is not, for instance, when neither of the two power stations operate, the generator in b is out of order, yet house a is supplied with electricity by its own generator L. Accordingly, L is a determinant of C but not of E. In case of interpretation (II), hence, expanding $\mathcal{F}_1$ supplements our data in table 1 by additional coincidences to the effect that the entanglement of C and E is suspended. This yields a coincidence list that is unambiguously assignable to an (ordinary) epiphenomenon.

This difference with respect to expansions of chains and epiphenomena can be nicely illustrated by graphical means. Graph (d*) of figure 7 represents a possible
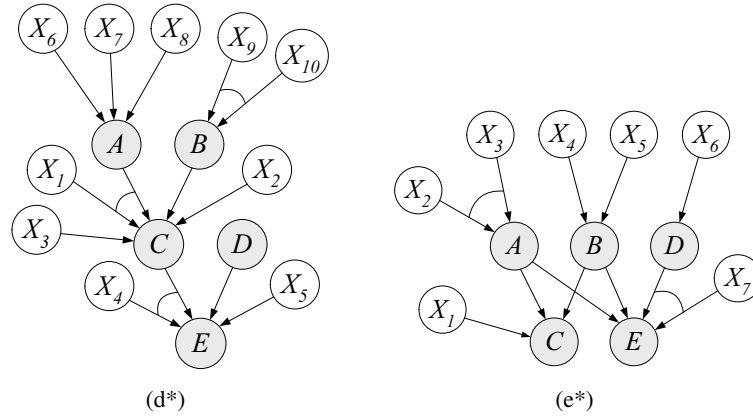
Fig. 7: Two possible expansions of graphs (d) and (e) in figure 2. Every determinant of C introduced into (d) above the C vertex, necessarily, is a determinant of E. The same does not hold for (e).

expansion of (d) which demonstrates that every determinant of C which is newly introduced into (d) is a determinant of E as well. Take, for instance, X_6 in (d*). X_6 is part of a minimally sufficient condition of C, *viz.* X_1X_6. Accordingly, X_6 is moreover part of a minimally sufficient condition of E, *viz.* $X_1X_4X_6$. Contrary to the chain, an epiphenomenal structure does not necessitate newly introduced determinants of C to be determinants of E as well. X_1 in (e*) is a determinant of C but not one of E.

Of course, there again is an entangled epiphenomenon that is c-equivalent to (d*). Nonetheless, chains and entangled epiphenomena differ with respect to expansions of their factor frames. While chains *necessarily* conserve entanglements across all factor frame extensions, entangled epiphenomena *do not necessitate* the persistence of entanglements. The retention of entanglements across arbitrary factor frame or model extensions is a characteristic feature of deterministic chains. In every deterministic chain there are at least two entangled factors whose entanglement subsists across all extensions. That is, such entanglements are necessary conditions for a respective coincidence list to be the result of a causal chain. The same does not hold for epiphenomena. Even though certain factors may be entangled in epiphenomenal structures – as graph (e) demonstrates –, these structures do not necessitate such entanglements to be permanent.

Therefore, even though modeling a coincidence list as the one in table 1 in terms of a chain or an epiphenomenon is equally warranted by table 1, it makes a difference whether somebody causally analyzing that data takes it to be generated by a chain or by an epiphenomenon. In order to clearly bring out that difference let us picture two different experimenters investigating the structure behind table 1 – Fennella and Shamus. Assume that Fennella assigns chain (d) to table 1 and Shamus models that data in terms of epiphenomenon (e). Both Fennella and Shamus equally account for table 1. Fennella, however, not only provides a causal

model that adequately accounts for that data, but moreover assesses that no factor newly introduced into her model in the course of further analysis of the structure behind table 1 will resolve the entanglement of C and E. Put differently, Fennella does not only claim the causal dependencies to subsist that are stated in (d), but thereby also claims (CE). Accounting for the data in table 1 in terms of (d) implies (CE). In contrast, by assigning the entangled epiphenomenon (e) to table 1 Shamus merely states the dependencies expressed in (e). He remains completely non-committal with respect to (CE). (e) does not imply (CE). Hence, taking a co-incidence list as 1 to be generated by a chain is a *stronger claim* than modeling the causal structure behind table 1 in terms of an epiphenomenon.

Now, let us assume that Fennella and Shamus embark on further analysis of the causal structure behind table 1. That is, they extend the factor frame $\mathcal{F}_1$ and collect new coincidence data on the factors in the extended frame. Continuously broadening the empirical basis of their study in this vein, trivially, brings about either of two scenarios: The entanglement of C and E is not suspended – scenario $\mathcal{T}_1$ – or the entanglement of C and E is suspended – scenario $\mathcal{T}_2$. If the latter scenario obtains, i.e. if an extension of $\mathcal{F}_1$ leads to the discovery of a determinant of C which is not at the same time a determinant of E, the structure behind the behavior of the factors in $\mathcal{F}_1$ is proven to be epiphenomenal. Accordingly, Shamus's model is proven to be correct, whereas Fennella's is falsified. Suppose, however, further investigation into the causal structure behind $\mathcal{F}_1$ does not reveal a previously unknown factor that suspends the entanglement of C and E, i.e. suppose $\mathcal{T}_1$ obtains. Both Fennella and Shamus can extend their models such that they account for scenario $\mathcal{T}_1$. A continuing entanglement of C and E can be modeled either in terms of a chain or of an epiphenomenon. However, Fennella predicted that $\mathcal{T}_1$ would occur from the beginning, whereas Shamus did not commit himself to either $\mathcal{T}_1$ or $\mathcal{T}_2$. Thus, finding that $\mathcal{T}_1$ in fact obtains does not surprise Fennella at all. Fennella's claim (CE) receives empirical support by $\mathcal{T}_1$. Of course, the mere fact that (CE), which is a logical consequence of Fennella's chain model, is empirically confirmed by $\mathcal{T}_1$ does not prove the chain model to be correct. Scenario $\mathcal{T}_1$ lends equal empirical support to the chain as to the common cause model. Nonetheless, Fennella can straightforwardly account for why $\mathcal{T}_1$ *rather than* $\mathcal{T}_2$ has occurred. In contrast, even though Shamus can suitably extend his original model (e) such that every newly introduced determinant of C is also a determinant of E, his epiphenomenal model by itself does not account for why $\mathcal{T}_1$ rather than $\mathcal{T}_2$ has occurred. Maybe Shamus can provide a different reason as to why the extension of the original frame $\mathcal{F}_1$ did not lead to the discovery of a determinant of C that is not a determinant of E as well – maybe prior causal knowledge about the process under investigation is available that accounts for why $\mathcal{T}_1$ rather than $\mathcal{T}_2$ has occurred. Without such an independent account for the subsistence of the entanglement of C and E, however, finding that this entanglement has not been broken by further data must be somewhat surprising to Shamus. The epiphenomenal model (e) does not provide any reason to expect $\mathcal{T}_1$ rather than $\mathcal{T}_2$ to occur.

That means a chain and an associated entangled epiphenomenon have different implications for expansions of the empirical basis of a causal analysis. This asymmetry might be fruitfully exploited when it comes to choosing between the two models. The previous section has casted doubt on there existing purely empirical reasons that unambiguously call for modeling deterministic data that features at least two entangled factors in terms of a chain. The methodology of causal reasoning that fares best when it comes to inferring deterministic chains arguably is one embedded in an interventionist framework as e.g. propagated by Woodward (2003). Yet, we have seen that interventionist methodologies presuppose a causally entrenched notion of an intervention that excludes one of two possible models that equally account for the entanglement of two factors by assumption. The different implications of the chain and the epiphenomenal model with respect to factor frame expansions could now be taken as grounds on which such disambiguating causal assumptions can be justified. Fennella's chain model implies that scenario $\mathcal{T}_1$ rather than $\mathcal{T}_2$ obtains and, thus, can be said to account for the fact that (CE) holds in our exemplary case. As long as Shamus cannot come up with an account of why $\mathcal{T}_1$ rather than $\mathcal{T}_2$ has occurred that fares equally well as Fennella's, a causal assumption that excludes the epiphenomenal model in the vein of the interventionist framework is well-grounded and justified.

7 Conclusion

This paper has answered one question and raised another. It has negatively answered the question as to whether ambiguities of causal reasoning are a peculiarity of causally analyzing (noisy) probabilistic data. Analyzing deterministic structures against ideally homogeneous backgrounds such that deterministic dependencies are exhibited in the data by no means guarantees unambiguous causal inferences. We have found that inferences to causal chains are systematically empirically underdetermined. Whenever deterministic data could stem from a chain, it could equally be the result of a common cause structure. The finding that deterministic chains cannot be positively identified on purely empirical grounds raises the question on what other grounds inferences to deterministic chains could be justified. We have seen that even though chain and associated common cause models are empirically equivalent, they have different implications for expansions of the empirical basis of a pertaining study. These differences could be resorted to in order to justify causal assumptions that allow for a positive identification of causal chains. However, this paper has at best provided some hints at how the inference to deterministic chains could be non-empirically disambiguated. A general account of how causal chains can be conclusively assigned to noise-free deterministic data has to await another paper.

References

Baumgartner, Michael: *Complex Causal Structures. Extensions of a Regularity Theory of Causation*, Ph. D thesis, University of Bern 2006.

—— Regularity Theories Reassessed, *Philosophia* ((2008)).

—— Uncovering Deterministic Causal Structures: A Boolean Approach, *Synthese* (2008).

—— Inferring Causal Complexity, (unpublished).

Brand, Myles: Simultaneous Causation, in: Inwangen, Peter van, editor: *Time and Cause*, Dordrecht: Reidel 1980, 137–153.

Brayton, Robert K. and Khatri, Sunil P.: Multi-Valued Logic Synthesis, in: *VLSID '99: Proceedings of the 12th International Conference on VLSI Design - 'VLSI for the Information Appliance'*, Washington, DC, USA: IEEE Computer Society 1999, 196.

Dowe, Phil: *Physical Causation*, Cambridge: Cambridge University Press 2000.

—— Causal Processes, in: Zalta, Edward N., editor: *The Stanford Encyclopedia of Philosophy*, Fall 2007 ⟨URL: http://plato.stanford.edu/archives/fall2007/entries/causation-process/⟩.

Frydenberg, Morten: The Chain Graph Markov Property, *Scandinavian Journal of Statistics* 17 (1990), 333–353.

George, Alexander L. and Bennett, Andrew: *Case Studies and Theory Development in the Social Sciences*, Cambridge: MIT Press 2005.

Glymour, Clark: A Review of Recent Work on the Foundations of Causal Inference, in: McKim, Vaughn R. and Turner, Stephen P., editors: *Causality in Crisis?* Notre Dame: University of Notre Dame Press 1997, 201–248.

—— Learning the Structure of Deterministic Systems, in: Gopnick, A. and Schulz, L., editors: *Causal Learning. Psychology, Philosophy, and Computation*, New York: Oxford University Press 2007, 231–240.

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

Huemer, Michael and Kovitz, Ben: Causation as Simultaneous and Continuous, *The Philosophical Quarterly* 53 (2003), 556–565.

Kistler, Max: Causation as Transference and Responsibility, in: Spohn, W., Ledwig, M., and Esfeld, M., editors: *Current Issues in Causation*, Paderborn: Mentis 2001, 115–133.

Lewis, David: Counterfactual Dependence and Time's Arrow, *Noûs* 13 (1979), 455–476.

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

Mahoney, James: Strategies of Causal Inference in Small-N Analysis, *Sociological Methods & Research* (2000), 387–424.

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

Pearl, Judea: Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning, in: *Proceedings of the 7th Conference of the Cognitive Science Society*, Irvine: University of California 1985.

—— *Causality. Models, Reasoning, and Inference*, Cambridge: Cambridge University Press 2000.

Quine, Willard van Orman: The Problem of Simplifying Truth Functions, *The American Mathematical Monthly* 59 (1952), 521–531.

——— On Core and Prime Implicants of Truth Functions, *American Mathematical Monthly* 66 (1959), 755–760.

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

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

Richardson, Thomas, Schulz, Laura, and Gopnik, Alison: Data-Mining Probabilists or Experimental Determinists? in: Gopnik, Alison and Schulz, Laura, editors: *Causal Learning. Psychology, Philosophy, and Computation*, Oxford: Oxford University Press 2007, 208–230.

Salmon, Wesley: Causality Without Counterfactuals, *Philosophy of Science* 61 (1994), 297–312.

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

Steinmetz, George: Critical Realism and Historical Sociology, *Comparative Studies in Society and History* 40 (1998), 170–186.

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

Verma, T. S. and Pearl, Judea: Equivalence and Synthesis of Causal Models, in: Bonissone, Piero P. et al., editors: *Uncertainty in Artificial Intelligence 6*, Amsterdam: Elsevier 1991, 255–268.

Woodward, James: *Making Things Happen*, Oxford: Oxford University Press 2003.