Causality, Conditional Independence, and Graphical Separation in Settable Systems

Authors: Karim Chalak, Halbert White

This work is posted on eScholarship@BC, Boston College University Libraries.


Originally posted on: http://ideas.repec.org/p/boc/bocoec/689.html
Causality, Conditional Independence, and Graphical Separation in Settable Systems

Karim Chalak*† Halbert White  
Boston College UC San Diego

July 4, 2010

Abstract

We study the connections between conditional independence and causal relations within the Settable Systems extension of the Pearl Causal Model. Our analysis clearly distinguishes between causal notions and probabilistic notions and does not formally rely on graphical representations. We provide definitions in terms of functional dependence for direct, indirect, and total causality as well as for indirect causality via and exclusive of a set of variables. We then provide necessary and sufficient causal and probabilistic conditions for conditional dependence among random vectors of interest in structural systems. We state and prove the conditional Reichenbach principle of common cause, obtaining the classical Reichenbach principle as a corollary. Finally, we relate our results to notions of graphical separation such as $d$-separation and $D$-separation in the artificial intelligence and machine learning literature.

Running Title: Conditional Independence in Causal Systems  
Keywords: causality, conditional independence, $d$-separation, Reichenbach principle, settable systems.

1 Introduction

In the last two decades, graphical representations of probabilistic relations, and in particular conditional independence relations (see e.g. Dawid, 1979, 1980), have been extensively

*Karim Chalak is Assistant Professor of Economics, Dept. of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467 (email: chalak@bc.edu); and Halbert White is Chancellor’s Associates Distinguished Professor of Economics, Dept. of Economics 0508, UCSD, La Jolla, CA 92093 (email: hwhite@ucsd.edu).

†An earlier version of this paper was circulated under the title “Independence and Conditional Independence in Causal Systems.” We thank participants of the Harvard/MIT, Yale, NYU, and Pennsylvania State University Econometrics Seminars, Julian Betts, Graham Elliott, Clive Granger, Arthur Lewbel, Mark Machina, Dimitris Politis, Sin-Miao Wang, and especially Ruth Williams for their helpful comments and suggestions. Any errors are solely the authors’ responsibility.
studied in machine learning and statistics (e.g. Lauritzen and Spiegelhalter, 1988; Pearl, 1988, 2000; Lauritzen, Dawid, Larsen, and Leimer, 1990; Geiger, Verma, and Pearl, 1990; Lauritzen and Richardson, 2002; Wermuth and Cox, 2004). A particularly fruitful approach uses directed acyclic graphs (DAGs) to represent a collection of conditional independence relations (e.g. Lauritzen et al., 1990; Geiger et al., 1990). DAGs are also widely used to represent and even to define causal relations, as in Spirtes, Glymour, and Scheines (1993, hereafter SGS) and the Pearl Causal Model (PCM) (Pearl, 2000, pp. 202-205).

The use of a DAG both to represent probabilistic relations among random variables on the one hand and causal relations among variables on the other has been exploited to infer interrelations between probabilistic and causal relations, leading to significant developments in machine learning and causal inference (see e.g. SGS; Pearl, 1993, 1995, 2000; Shpitser and Pearl, 2008). At the center of much of this interplay between association and causation is Reichenbach’s (1956) “principle of common cause,” which states that if two variables are associated (e.g., correlated) then either one “causes” the other or they both share a third “common cause.” Although this principle has intuitive appeal and despite its venerated status, its formal standing so far remains ambiguous (see e.g. Dawid, 2010a, p. 66).

Dawid (2002, 2010a, 2010b) warns that in order to represent causal concepts such as “direct causal effect” using a DAG, these must be defined a priori “by other, necessarily non-graphical, considerations not involving these terms” (Dawid, 2010, p. 66). The PCM is a framework within which certain causal and probabilistic concepts have been defined and represented by a DAG. However, Dawid (2002, 2010a, 2010b) argues that a clear separation of causal and probabilistic semantics and an explicit statement of the imposed assumptions is needed to justify the simultaneous use of a DAG to appropriately represent probabilistic and causal relations embodied in the PCM. Dawid (2002, 2010a, 2010b) advocates the use of “extended conditional independence” relations and their representations using “influence diagrams” to achieve this. Similar concerns regarding the appropriateness of the assumptions underlying DAGs are raised in Duvaud, Eaton, Murphy, and Schmidt (2010), leading them to adopt a “black box” view of causal models in which causality is defined in functional terms but where causal models are evaluated in terms of their predictive performance. Like Duvaud et. al., we favor a function-based definition of causality, although we concur with Pearl (2000, p. 61) that “fitness to data is an insufficient criterion for validating causal theories,” as non-causal relations can easily deliver superior predictions. We also completely agree with Dawid (2010a) that a rigorous treatment is warranted that unambiguously distinguishes between notions of causality and of probabilistic dependence and that also makes a clear separation between these notions and their
graphical representations (if any). Although Dawid (2002, 2010a, 2010b) presents a series of instructive examples and discussions, these papers nevertheless do not put forward a self-contained formal framework accomplishing this. Nor does this exist elsewhere.

A main goal of this paper, therefore, is to provide such a framework, by formally studying the connections between conditional independence relations and causal relations within the framework of \textit{settale systems} proposed by White and Chalak (2009) (henceforth “WC”) as an extension of the PCM. Rather than building on the properties of probabilistic DAGs, we begin with definitions of causality based on functional dependence, defined within a given settable system. From these naturally emerge graphical representations (directed graphs and DAGs) that are helpful to heuristic understanding. While helpful for intuition, graphs do not play a formal role in our analysis. We then study how the presence or absence of causal relations in a given settable system give rise to specific independence and conditional independence relations. We relate our results to criticisms and suggestions proposed in Dawid (2002, 2010a, 2010b) to demonstrate how the settable system framework permits addressing these, while preserving many of the appealing features of the PCM.

Our results thus shed light on two fundamental questions. First, what implications for their joint probability distribution derive from knowledge of functionally defined causal relationships between variables of interest? Conversely, what restrictions (if any) on the possible functionally defined causal relationships holding between variables of interest follow from knowledge of the probability distribution governing these variables? This contributes to the understanding of empirical relationships by elucidating the linkage between causal relations known or theorized to underlie a body of data and the joint probability distribution of the data, especially as reflected in conditional independence relations.

This paper is organized as follows. In Section 2, we briefly discuss a number of related strands of the literature that provide background and motivate several specific further contributions of the present paper, e.g., our extension of prior notions of indirect causality and our proof of the conditional Reichenbach principle of common cause. Section 3 provides further motivation by illustrating certain problems and limitations, similar to those discussed in Dawid (2002, 2010a, 2010b), that arise when studying the connections between probabilistic and causal relations using probabilistic DAGs, the PCM, and PCM DAGs. For this, we use an example in which an expert advises an agent on an action influencing an outcome of interest. Section 4 revisits this example, placing it in the settable systems framework and showing how this overcomes limitations of the PCM discussed in Section 3. Sections 5, 6, 7, and 8 formalize and extend Section 4’s material. Section 5 formally introduces a version of WC’s settable systems conveniently suited to formulating rigorous
definitions, provided in Section 6, of direct causality based on functional dependence, as well as notions of indirect causality via a set of variables and exclusive of a set of variables in recursive systems. Sections 7 and 8 provide connections between causality and conditional independence. Section 7 introduces and proves the conditional Reichenbach principle of common cause. We then provide necessary and sufficient conditions for probabilistic conditional dependence of certain vectors of random variables in recursive settable systems. The traditional Reichenbach principle obtains as a corollary. In Section 8, we relate our results to $d$-separation and $D$-separation (discussed in Geiger et. al., 1990), and we study properties of restricted settable systems analogous to Markovian and semi-Markovian PCMs. In particular, we provide conditions sufficient for causal intuitions attributed to $d$—separation or $D$—separation to hold in recursive settable systems. Section 9 concludes and discusses directions for future research. Formal mathematical proofs are collected in the Mathematical Appendix.

2 Background and Relation to the Literature

2.1 Probabilistic DAGs

An important contribution of the artificial intelligence literature is the introduction of graphical criteria applicable to DAGs that characterize independence and conditional independence relations among variables in “Bayesian Networks” or, more specifically, “directed Markov fields.” In these DAGs, each node represents a random variable. For example, in graph $G_1$ we have 5 random variables $X_1, ..., X_5$. A DAG is said to represent a probability distribution for these random variables when the joint density function exists and factorizes as the product of the densities of each random variable conditional on its “parents” in the graph. For example, in $G_1$ we have:

$$p(x_1, x_2, x_3, x_4, x_5) = p_1(x_1)p_2(x_2)p_3(x_3|x_1, x_2)p_4(x_4|x_2)p_5(x_5|x_4),$$

where the left-hand term denotes the joint density and each right-hand term denotes the density of one variable conditional on the value of its “parents.” Following Dawid (2002),
we refer to DAGs of this kind as “probabilistic DAGs.”

Lauritzen et al. (1990, theorem 1) show that the joint density admits such a recursive factorization if and only if the collection of conditional independence statements that each variable is conditionally independent of its “non-descendants” given its “parents” in the DAG holds. Lauritzen et al. (1990) refer to the latter property as the “directed local Markov property”; SGS (p. 54) refer to this as the “causal Markov property”; and Pearl (2000, theorem 1.2.7) calls it the “parental Markov condition.” Using Dawid’s (1979) notation \( \perp \) to denote independence, \( G_1 \) implies, for example, that \( X_1 \perp X_2 \) and \( X_3 \perp X_4 \mid (X_1, X_2) \) for any distribution represented by \( G_1 \).

### 2.2 Attributing Causal Meaning to Probabilistic DAGs

A causal meaning is sometimes attributed to such DAGs. In particular, a directed arrow from \( X_2 \) to \( X_3 \) in \( G_1 \) is interpreted to mean that “\( X_2 \) is a direct cause of \( X_3 \).” But there is no formal basis whatsoever for such interpretations in probabilistic DAGs. As Dawid (2002, p. 164) states, “there is absolutely nothing in the probabilistic semantics by which such graphs are supposed to be interpreted that is relevant to such causal intuitions.”

Pearl (2000, definitions 3.2.1 and 4.5.1), however, provides definitions for “total” and “direct” causal effects within the PCM, linking these concepts to the connectivity properties of corresponding DAGs. Pearl (2001) and Avin, Shpitser, and Pearl (2005) similarly provide definitions for "indirect" effects as well as “path-specific” effects. Related notions of direct, indirect, and total effects have been proposed in Robins and Greenland (1992), SGS, Robins (2003), Didelez, Dawid, and Geneletti (2006), and Geneletti (2007); see also Rubin (2004).

We resolve the apparent contradiction between the causal use of DAGs in the PCM and Dawid’s cogent warnings by providing rigorous definitions of direct and indirect causality based on functional dependence. These functional relations do not depend on graphs, but they do lend themselves to convenient graphical representation. Our definitions extend previous notions of indirect causality to accommodate notions of causality via a set of variables
and *exclusive of* a set of variables in recursive systems. Although these extensions are of interest in their own right, their larger significance is that they provide suitable foundations for rigorously proving and extending Reichenbach’s (1956) “principle of common cause.”

Reichenbach’s principle is central to certain strands of the philosophy literature (e.g., Spohn, 1980; Hausman and Woodward, 1999; Cartwright, 2000) that examine the connections between causal structure and conditional independence relations. Reichenbach’s principle states that if two variables are associated (e.g., correlated) then either one “causes” the other or they both share a third “common cause.” For example, the assumption $X_1 \perp X_2$ in $G_1$ may be attributed to the lack of a common cause of $X_1$ and $X_2$. Despite its intuitive appeal and venerated status, this principle’s formal standing is nevertheless ambiguous (see e.g. Dawid, 2010a, p. 66). Is it an axiom or a postulate, or is it a logical consequence of assumptions as yet unformulated?

Another contribution of this paper is to provide a formal answer to this question. Specifically, we show that Reichenbach’s principle follows as a logical consequence of the assumptions defining settable systems. In fact, Reichenbach’s principle follows as a corollary to a more general result that we call the *conditional* Reichenbach principle of common cause. This result leads to necessary and sufficient conditions for probabilistic conditional dependence of certain vectors of random variables in settable systems. As immediate corollaries, we obtain straightforward causal conditions sufficient to ensure or rule out independence or conditional independence among random vectors in settable systems.

### 2.3 $d$-Separation

Using properties of conditional independence (e.g. Dawid, 1979; Studeny, 1993), one can infer further conditional independence statements that hold among the variables represented in a probabilistic DAG. In particular, Geiger, et. al. (1990) (see also Verma and Pearl, 1988; Geiger and Pearl, 1993; Pearl, 2000) provide a graphical criterion, called “$d$-separation,” that can identify exactly the conditional independence relations implied by a probabilistic DAG under the “graphoid” axioms\(^1\). Lauritzen et. al. (1990, proposition 3) provide a graphical criterion equivalent to $d$-separation and show that the implications of these criteria when applied to a probabilistic DAG are equivalent to the directed local Markov property (Lauritzen et. al., 1990, theorem 1).

For example, in probabilistic DAG $G_1$, one can inspect whether nodes $X_i$ and $X_j$ are $d$-separated by a set of nodes $W \subseteq \{X_1, \ldots, X_5\}/\{X_i, X_j\}$. For this, let an $(X_i, X_j) − \text{trail}$

---

\(^1\)In Geiger, et. al. (1990), the four graphoid axioms are properties of conditional independence relations discussed, for example, in Dawid (1979).
in $G_1$ be any sequence of arrows linking $X_i$ to $X_j$ irrespective of their directionality. Then $W$ $d-$separates $X_i$ and $X_j$ in $G_1$ if every $(X_i, X_j)$-trail in $G_1$ contains either (1) a node $W_k \in W$ that does not have converging arrows along the $(X_i, X_j)$-trail, or (2) a node $X_k$ that has converging arrows along the $(X_i, X_j)$-trail, such that neither $X_k$ nor any of its descendants are in $W$ (Geiger et al., 1993; Pearl 2000, definition 1.2.3; see also Lauritzen et. al., 1990 for an equivalent graphical criterion). For example, one can conclude from $G_1$ that $X_3 \perp X_4 | X_2$ since $X_3$ and $X_4$ are $d$-separated by $X_2$, and that $X_2 \perp X_5 | X_4$ since $X_4$ $d$-separates $X_2$ and $X_5$.

2.4 Attributing Causal Meaning to $d$-Separation

Implications of $d$-separation have been ascribed causal intuition (see for e.g. Pearl, 2000, p. 16-17). In example $G_1$, $d$-separation implies $X_2 \perp X_5 | X_4$, which has been interpreted to mean that conditioning on a variable $X_4$ that fully mediates the effect of a cause $X_2$ on a response $X_5$ renders $X_2$ and $X_5$ conditionally independent. Similarly, $X_3 \perp X_4 | X_2$ has been interpreted to mean that conditioning on the common cause $X_2$ of the two effects $X_3$ and $X_4$ renders $X_3$ and $X_4$ conditionally independent. Also, the fact that $X_1 \perp X_2 | X_3$ is not implied by $d$-separation has been attributed to the notion that conditioning on a common response $X_3$ of causes $X_1$ and $X_2$ renders these conditionally dependent.

We emphasize that, consistent with Dawid’s warnings, there is no formal basis for such causal interpretations in probabilistic DAGs. Our final contribution is to show that in our settable systems framework, such causal statements are fully meaningful. That is, causal statements motivated by graphical intuitions make sufficient sense that they can be determined to be true or false. As we discuss, $d$-separation in probabilistic DAGs is neither necessary nor sufficient for the presence or absence of specific causal effects. Nevertheless, we describe certain restricted settable systems in which causal interpretations hold that are similar to the causal intuitions derived from $d$-separation.

3 A Motivating Example

To motivate the formal results presented in the subsequent sections, we consider a simple scenario where an expert $e$ advises an agent $a$ on an action that may influence an outcome of interest to $a$. For example, $e$ may be a physician recommending a medical treatment to patient $a$, or $e$ may be a financial expert recommending an investment plan to investor $a$. 

7
3.1 A Probabilistic DAG

Consider the following simple probabilistic DAG involving variables $Y_1^c$, $Y_2^c$, and $Y_3^c$ denoting the advice of expert $e$, action of agent $a$, and outcome, respectively.

![Graph 2 (G2)](image)

Probabilistically, graph $G_2$ shows that $e$’s advice is independent of the outcome given $a$’s action, $Y_1^c \perp Y_3^c \mid Y_2^c$. This follows due to the lack of an arrow between $Y_1^c$ and $Y_3^c$, implying $d$-separation.

3.2 A Pearl Causal Model

A PCM (Pearl, 2000, definition 7.1.1) for this example assumes that each of the “endogenous” variables ($Y_1^c$ for advice, $Y_2^c$ for action, and $Y_3^c$ for outcome), is determined as a function of its “parents” and “background variables” that are “often unobservable” (Pearl, 2000, p. 203). For simplicity, let $U_1$, $U_2$, and $U_3$ be random background variables each associated with the endogenous variables $Y_1^c$, $Y_2^c$, and $Y_3^c$ respectively as in Pearl (2000, p. 68), for example. In particular, suppose that

$$
Y_1^c = g_1(U_1),
Y_2^c = g_2(Y_1^c, U_2), \text{ and }
Y_3^c = g_3(Y_2^c, U_3),
$$

with $g_1$, $g_2$, and $g_3$ denoting “potential response” functions. Observe that we assume that $Y_1^c$ is not an explicit argument of the potential response function $g_3$.

As discussed in WC, the PCM rules out any causal role for the background $U$’s, since these are not subject to “counterfactual variation” (see Pearl 2000, definition 7.1.3). Since $g_3$ excludes $Y_1^c$ from its arguments, we assume that $e$’s advice does not “directly cause” the outcome. Also implicit in the above PCM is the assumed exclusion of endogenous variables other than $Y_1^c$, $Y_2^c$, and $Y_3^c$. Dawid (2010b) argues for explicitly referencing the “causal ambit,” that is, the set of variables which the subset of endogenous variables $\{Y_1^c, Y_2^c, Y_3^c\}$ belongs to, in order to discuss notions such as unobserved “common cause,” etc.; we return to this shortly.

---

2The superscript “c” denotes “canonical” variables arising from the “natural” (unmanipulated) operation of the system. This conforms with notation formally introduced below.
3.3 \textit{d}-Separation and Conditional Independence in PCM DAGs

How, if at all, can this PCM generate conditional independence relations among $Y_1^c$, $Y_2^c$, and $Y_3^c$ that coincide with those encoded by the probabilistic DAG $G_2$? To answer this question, suppose that background variables $U_1$, $U_2$, and $U_3$ corresponding to the advice, action, and outcome are jointly independent. This assumption yields a “Markovian model,” in which the jointly independent “arbitrary distributed random disturbances [...] represent independent background factors that the investigator chooses not to include in the analysis.” (Pearl, 2000, pp. 68 - 69).

Now consider the PCM DAG associated with this Markovian model. In PCM DAGs, arrows between endogenous variables denote “direct causal relations” (e.g. Pearl, 2000). Thus, the assumption that $e$’s advice does not “directly cause” the outcome is represented by a missing arrow from $Y_1^c$ to $Y_2^c$. Typically, only the endogenous variables $Y_1^c$, $Y_2^c$, and $Y_3^c$ are represented at the nodes of a Markovian PCM DAG (e.g. Pearl, 2000, chapters 3 and 5). This yields the PCM DAG depicted by $G_2$ which is isomorphic (has identical connectivity) to the probabilistic DAG also depicted by $G_2$.

Using this PCM structure and properties of conditional independence relations, it can be shown\textsuperscript{3} that $Y_1^c \perp Y_3^c \mid Y_2^c$, as represented in probabilistic DAG $G_2$. In this case, the PCM represented by the PCM DAG $G_2$ generates conditional independence relations among the endogenous variables $Y_1^c$, $Y_2^c$, and $Y_3^c$ that coincide with the conditional independence relations encoded via the \textit{d}-separation criterion in probabilistic DAG $G_2$ isomorphic to PCM DAG $G_2$.

3.4 Conditional Independence without \textit{d}-Separation

Suppose now that agent $a$ fully complies with expert $e$’s advice. One way to represent this is to exclude background variable $U_2$ from the arguments of $g_2$ so that $Y_2^c = g_2(Y_1^c)$. In this case, arguments similar to those in Section 3.3 give that outcome $Y_3^c$ is independent of advice $Y_1^c$ given action $Y_2^c$. But now we also have that action $Y_2^c$ is independent of outcome $Y_3^c$ given advice $Y_1^c$, that is $Y_2^c \perp Y_3^c \mid Y_1^c$, despite the fact that $Y_2^c$ and $Y_3^c$ are not \textit{d}-separated by $Y_1^c$ in probabilistic DAG $G_2$. Here, the PCM generates conditional independence relations that are not encoded via \textit{d}-separation in probabilistic DAG $G_2$.

Because $Y_2^c$ is fully determined by $Y_1^c$, Geiger, et. al. (1990) refer to $Y_2^c$ as a “deter-
“

3.5 \textit{d}-Separation without Conditional Independence

Now suppose that the background variables $U_1$, $U_2$, and $U_3$ are not mutually independent. Because we do not modify the causal relations among the endogenous variables, the arrows linking endogenous variables in PCM DAG $G_2$ remain unaltered. However, the arguments from Section 3.3 to establish conditional independence of the advice and outcome given the action are no longer valid and $Y_1^c \not\independent Y_3^c \mid Y_2^c$ may hold, despite the fact that $Y_1^c$ and $Y_3^c$ are $d$-separated by $Y_2^c$ in PCM DAG $G_2$. One can choose a particular distribution for the background variables and particular potential response functions to ensure the conditional independence of the advice $Y_1^c$ and outcome $Y_3^c$ given the action $Y_2^c$. Nevertheless, the functions and underlying probability distribution needed for this are extremely special.

In the absence of such special conditions, $Y_1^c \not\independent Y_3^c \mid Y_2^c$ generally, so the PCM does not generate conditional independence relations that are implied by $d$-separation in the probabilistic DAG corresponding to the PCM DAG $G_2$. Here, too, the PCM DAG must be augmented with quite special structure to guard against incorrect inference based on $d$-separation about probabilistic relations holding among the PCM’s endogenous variables. To show dependence, the PCM DAG $G_2$ is augmented with bidirected arcs between nodes corresponding to endogenous variables whose background variables are not independent.

Since the PCM rules out “exogenous” causes, one way to accommodate these is to assign certain background variables a causal status within the PCM so that they then become endogenous. Indeed, to facilitate applying the $d$-separation criterion to PCM DAGs, a bidirected arc linking two endogenous variables is often replaced in the PCM DAG by an unobserved endogenous “common cause” of the two endogenous variables (see e.g. Pearl 2000, chapters 3 and 5). But this implies that to study the connections between conditional independence relations and causal relations within the PCM framework, one must: (1) specify the “causal ambit” discussed in Dawid (2010b); (2) specify which observables and unobservables are endogenous; and, importantly, (3) assume independence or dependence
relations among background variables that do not have a causal status.

The existence of these jointly independent background variables is a strong assumption. Often, they do not emerge naturally from the system of interest; rather, they appear artificial. As Dawid (2002, p. 183) observes, “when the additional variables are pure mathematical fictions, introduced merely so as to reproduce the desired probabilistic structure of the domain variables, there seems absolutely no good reason to include them in the model.”

4 Settable Systems Formulation

In this section, we place the example above within the settable system framework. This permits a clear separation between causal and probabilistic semantics, thereby addressing the difficulties arising for the PCM in connecting causal and probabilistic relations discussed in Section 3.

A general model for our advice, action, and outcome scenario should represent an uncertain environment in which a’s action may respond to a setting of e’s recommendation. Throughout, all random variables are defined on a measurable space $(\Omega, \mathcal{F})$. We let the space $\Omega$ be sufficiently rich so that components of $\Omega := \times_{i=0}^{n} \Omega_i = \times_{i=0}^{3} \Omega_0$, with each $\Omega$, a copy of the principal space $\Omega_0$, underlie settings and responses, as we now discuss. A setting of e’s recommendation is a random variable $Z_1 : \Omega_1 \to S_1$ with support $S_1 \subseteq \mathbb{R}$. Because agent a may or may not implement e’s recommendation, the model does not assume that e’s recommendation entirely determines a’s action. A general way to capture this sort of dependence is to express the response $Y_2$ of a’s action by

$$Y_2(\omega) := r_2(\omega_0, Z_1(\omega_1)),$$

where $r_2$ is a real-valued measurable response function, and $\omega := (\omega_0, ..., \omega_3) \in \Omega$ represents a “possibility,” with components $\omega_1$ determining the value of setting $Z_1$, and $\omega_0$, a “state of nature,” determining the response value $Y_2(\omega)$ given the setting value $Z_1(\omega_1)$. The response function $r_2$ can be determined by a governing principle, such as optimization. In this case, it represents the action that is best for a in some sense given the recommendation setting $Z_1$, and under uncertainty.

Similarly, a setting $Z_2 : \Omega_2 \to S_2$ of a’s action and a setting $Z_1$ of e’s recommendation may influence the response of a’s outcome, denoted $Y_3$, so that

$$Y_3(\omega) := r_3(\omega_0, Z_1(\omega_1), Z_2(\omega_2)).$$
Observe that unlike the PCM in Section 3.2, the setting $Z_2$ of $a$’s action need not coincide with the response $Y_2$ of $a$’s action. In settable systems, these objects are distinct.

The PCM permits degenerate settings via submodels (Pearl, 2000, definition 7.1.2), given a unique fixed point assumption (see WC); but it does not accommodate random settings, such as $Z_2$. In non-experimental situations, we may conceive of a setting $Z_2$ of $a$’s action even if it is not implemented.

In settable systems, both the setting $Z_2$ and response $Y_2$ refer to $a$’s action. Together, they define the settable variable $X_2 : \{0,1\} \times \Omega \rightarrow S_2$ for $a$’s action by

$$X_2(0,\omega) := Y_2(\omega) \quad \text{and} \quad X_2(1,\omega) := Z_2(\omega), \quad \omega \in \Omega.$$  

Similarly, the outcome setting $Z_3 : \Omega_3 \rightarrow S_3$ and response $Y_3$ define the settable variable $X_3$ for $a$’s outcome given by $X_3(0,\omega) := Y_3(\omega)$ and $X_3(1,\omega) := Z_3(\omega)$.

Observe that in this system, the expert’s recommendation is not influenced by settings $Z_2$ in individual $a$’s action or $Z_3$ in $a$’s outcome. Instead, the response $Y_1$ for $e$’s recommendation is given by $Y_1(\omega) := r_1(\omega_0)$. This endows the settable variable for $e$’s recommendation given by $X_1(0,\omega) := Y_1(\omega)$ and $X_1(1,\omega) := Z_1(\omega_1)$ with the distinctive feature that its response does not depend on any setting in the system other than its direct dependence on $\omega_0$. For this reason, we call $X_1$ a fundamental settable variable.

Last, it is useful to define the principal setting $Z_0$ as the identity mapping $Z_0 : \Omega_0 \rightarrow \Omega_0$ and the corresponding principal response $Y_0$ and principal settable variable $X_0$ such that $Z_0(\omega_0) := X_0(1,\omega) = \omega_0 = X_0(0,\omega) := Y_0(\omega)$. This obviates the need to introduce “background variables,” as in the PCM, to induce randomness in the responses. Instead, settable systems explicitly specify the dependence of the responses on other settings and on elements of the principal space $\Omega_0$, indexing “states of nature” (see also Heckerman and Shachter, 1995). Nor can we dispense with this structure without dispensing with the foundations needed to formalize stochastic behavior. Writing $\mathcal{X} := \{X_0, X_1, X_2, X_3\}$, we can represent this example as the settable system $\mathcal{S} := \{ (\Omega, F), \mathcal{X} \}$.

In $\mathcal{S}$, the responses $Y_1$ and $Y_2$ of $a$’s action and outcome are determined separately, as functions of all other system settings. Alternatively, we may consider what happens when $a$’s action and outcome responses are jointly determined under uncertainty, given a setting of $e$’s advice. To represent this, we partition the system’s $n = 3$ units into blocks. Specifically, consider blocking together units 2 and 3, separately from a block including just unit 1. This is represented by the partition $\Pi := \{\Pi_1, \Pi_2\}$, where $\Pi_1 = \{1\}$ and $\Pi_2 = \{2, 3\}$. In this case, responses $Y_i^\Pi$ and $Y_3^\Pi$ are jointly determined as

$$Y_i^\Pi(\omega) := r_i^\Pi(Z_1^\Pi(\omega_1), \omega_0), \quad \text{for } i \in \Pi_2,$$
with $Z_1^\Pi$ a setting of $e$’s advice under this particular partition. Thus, settings and response functions $r_i^\Pi$ are partition-specific. This system contrasts with $\mathcal{S}$, which is an elementary partitioned settable system. In the elementary partition, each unit $i$ forms its own block $\Pi_i^e = \{i\}$. For the remainder of this example, we work with the elementary partition $\Pi^e = \{\Pi_i^e, i = 1, 2, 3\}$ defining $\mathcal{S} = \mathcal{S}^e := \{([\Omega, F], (\Pi^e, \mathcal{X}^e))\}$, where now the partition and the partition dependence of the settable variables are made explicit.

An important feature of this particular system $\mathcal{S}^e$ is the inherent ordering of the variables such that settings of $\mathcal{X}_i$ may determine responses of $\mathcal{X}_j$ only if $i < j$. When this holds, we say that the system is a recursive partitioned settable system.

### 4.1 Causality in Settable Systems

What does it mean for $e$’s recommendation $\mathcal{X}_1$ to directly cause $a$’s action $\mathcal{X}_2$ in settable system $\mathcal{S}$? To formalize this notion, we first define an admissible intervention $(z_0, z_1, z_2, z_3) \rightarrow (z_0^*, z_1^*, z_2^*, z_3^*)$ to $(\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_2)$, to be two points $(z_0, z_1, z_2, z_3)$ and $(z_0^*, z_1^*, z_2^*, z_3^*)$ in the admissible space $\mathcal{S}_{[0:3]} \subseteq \Omega_0 \times \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3$, the space of all jointly admissible setting values. Underlying this intervention is a primary intervention $\omega \rightarrow \omega^*$, defined as two possibilities $\omega = (\omega_0, \omega_1, \omega_2, \omega_3)$ and $\omega^* = (\omega_0^*, \omega_1^*, \omega_2^*, \omega_3^*)$.

Often, we may hold constant all but one element in considering interventions, e.g., $\omega \rightarrow \omega^*$, where $\omega^* = (\omega_0, \omega_1^*, \omega_2, \omega_3)$, yielding setting values $z_1 = Z_1(\omega_1)$ and $z_1^* = Z_1(\omega_1^*)$ of $e$’s recommendations in a state of nature $z_0 = \omega_0$. Due to the recursive structure, any differences between $(\omega_2, \omega_3)$ and $(\omega_2^*, \omega_3^*)$ are irrelevant. Thus, it suffices here just to consider pairs of possibilities in $\mathcal{S}_{[0:1]} \subseteq \Omega_0 \times \mathcal{S}_1$. Generally, constraints on joint setting values $(z_0, z_1)$ may imply that $\mathcal{S}_{[0:1]} \neq \Omega_0 \times \mathcal{S}_1$; otherwise $\mathcal{S}_{[0:1]} = \Omega_0 \times \mathcal{S}_1$.

Causal relations in settable systems are defined as features of the response functions over their domain. For example, we say that $\mathcal{X}_1$ directly causes $\mathcal{X}_2$ in $\mathcal{S}$ if there exists an admissible intervention $(z_0, z_1) \rightarrow (z_0^*, z_1^*)$ such that $r_2(z_0, z_1^*) - r_2(z_0, z_1) \neq 0$, and we write $\mathcal{X}_1 \xrightarrow{D_S} \mathcal{X}_2$. Otherwise, we say $\mathcal{X}_1$ does not directly cause $\mathcal{X}_2$ in $\mathcal{S}$ and write $\mathcal{X}_1 \xrightarrow{\not D_S} \mathcal{X}_2$.

We emphasize that this defines causal relations in terms of settable variables, rather than in terms of arbitrary random variables or events. The latter have no necessary causal structure beyond that arising from the fact that random variables are measurable functions of some underlying $\omega$. In contrast, settable variables embody explicit structural relations holding among the variables of the system. This further entails that causal relations are always relative to a settable system, and, in particular, relative to the governing partition.

Similarly, we can formalize the notion of $\mathcal{X}_1$ directly causing $\mathcal{X}_3$ in system $\mathcal{S}$. Specifically,
a setting $Z_1$ of $e$’s recommendation may directly influence $a$’s outcome $Y_3$. For example, this may represent a form of “placebo effect” in the physician/patient example. We say that $X_1$ directly causes $X_3$ in $S$ if there exists an admissible intervention $(z_0, z_1, z_2) \rightarrow (z_0, z_1^*, z_2)$ to $(X_0, X_1, X_2)$ such that

$$r_3(z_0, z_1^*, z_2) - r_3(z_0, z_1, z_2) \neq 0,$$

and we then write $X_1 \xrightarrow{D} S X_3$. Otherwise, we say that $X_1$ does not directly cause $X_3$ in $S$, and we write $X_1 \not\xrightarrow{D} S X_3$. Below, we discuss the sense in which $X_1$ acts on $X_3$ without operating through any other system variable; in such cases we say that $X_1$ causes $X_3$ exclusive of $X_2$ in $S$, written $X_1 \xrightarrow{(2)} S X_3$. These two concepts are closely related, but it will be important to distinguish between them in what follows.

It can be useful to visually represent causal relations in $S$ using a direct causality graph. This consists of a collection of nodes corresponding to settable variables and a collection of directed arrows between nodes. A directed arrow links one node to another if and only if the first is a direct cause of the second. For example, graph $G_3$ visualizes possible direct causality relations in system $S$. We emphasize that direct causality graphs are neither probabilistic DAGs nor PCM DAGs.

### 4.2 Conditional Independence in Settable Systems

We now consider conditional independence relations among certain random variables in settable system $S$. Here, we focus on the responses of settable variables in “idle regimes” (Pearl, 2000; Dawid, 2010a), that is, in an observational environment where intervention or control are absent and the system evolves on its own. In particular, we focus on canonical settings of $e$’s recommendation and $a$’s action and outcome, denoted by $Z_1^c, Z_2^c,$ and $Z_3^c$, that are identical to responses to predecessor’s settings:\footnote{Note that $\omega_1, \omega_2,$ and $\omega_3$ appear as arguments of $Z_1^c, Z_2^c,$ and $Z_3^c,$ respectively. Thus, $\omega_1, \omega_2,$ and $\omega_3$ must be functions of $\omega_0$. Without loss of generality, we may take $\omega_1 = \omega_2 = \omega_3 = \omega_0.$}

\[
\begin{align*}
Z_1^c(\omega_1) &= Y_1^c(\omega_0) = r_1(\omega_0) \\
Z_2^c(\omega_2) &= Y_2^c(\omega_0) = r_2(\omega_0, Z_1^c(\omega_1)) \\
Z_3^c(\omega_3) &= Y_3^c(\omega_0) = r_3(\omega_0, Z_1^c(\omega_1), Z_2^c(\omega_2)).
\end{align*}
\]

Thus, given a possibility $\omega$, expert $e$ recommends $Y_1^c(\omega_0)$ based solely on the state of nature $z_0 = \omega_0$, and $a$ responds with $Y_2^c(\omega_0) = r_2(\omega_0, Y_1^c(\omega_0))$, yielding an outcome $Y_3^c(\omega_0) = r_3(\omega_0, Y_1^c(\omega_0), Y_2^c(\omega_0))$. We call $Y_1^c, Y_2^c,$ and $Y_3^c$ canonical responses. These correspond to
the PCM endogenous variables discussed in Section 3.2. We may also refer to the principal canonical setting and response $Z_0^c = Y_0^c = Z_0$.

Because we are interested in non-experimental environments, we focus here on conditional independence among canonical responses and not among arbitrary responses or settings, although these relations play a key role in other contexts. By focusing on canonical settings, we specify the “regimes” underlying canonical responses, similar to the discussion in Dawid (2002, 2010a, 2010b). Nevertheless, we maintain that response functions are “invariant” to different settings of a system’s settable variables. This is essentially without loss of generality, as the response function fully embodies the consequences for the response of any change in the argument settings.

Now let $P$ be a probability measure on $(\Omega, \mathcal{F})$. How can conditional independence relations hold among canonical responses $Y_1^c, Y_2^c,$ and $Y_3^c$? We distinguish between two possibilities. First, we consider conditional independence relations that hold among canonical responses for any probability measure. Second, we consider conditional independence relations that may hold only for some probability measures on $(\Omega, \mathcal{F})$.

To illustrate, suppose that $e$’s recommendation fully determines $a$’s action, such as when patient $a$ fully complies with doctor $e$’s advice. In this case, $\mathcal{X}_0$ has no impact on $\mathcal{X}_2$, except through $\mathcal{X}_1$. (Now that we are considering canonical systems, it is meaningful to speak of effects “through” or “via” other settable variables.) In this case, we have $\mathcal{X}_0 \sim^{(1)}_S \mathcal{X}_2$, so that for all admissible interventions to $(\mathcal{X}_0, \mathcal{X}_1)$ we have

$$r_2(z_0^*, z_1) - r_2(z_0, z_1) = 0,$$

Here, this also corresponds to $\mathcal{X}_0 \not\models_S \mathcal{X}_2$ (see graph $G_4$). Thus, we can write $Y_2(\omega) = r_2(\omega_0, Z_1(\omega_1)) = \tilde{r}_2(Z_1(\omega_1))$ for some measurable function $\tilde{r}_2$. In particular, canonical settings $Z_0^c$ and $Z_1^c$ yield the canonical response $Y_2^c = r_2(Y_0^c, Y_1^c) = \tilde{r}_2(Y_1^c)$. It follows that for any probability measure $P$, $\mathcal{X}_0 \not\models_S \mathcal{X}_2$ implies that $Y_2^c \perp Y_3^c \mid Y_1^c$. In this circumstance, we say that $\mathcal{X}_2$ and $\mathcal{X}_3$ are causally isolated given $\mathcal{X}_1$. The “isolation” is from the potential common cause $\mathcal{X}_0$. Because $G_4$ is a direct causality graph and not a probabilistic DAG, the notion of $d$-separation does not apply to it. Naively (mis)applying $d$-separation to $G_4$, we see that $\mathcal{X}_2$ and $\mathcal{X}_3$ are not $d$-separated by $\mathcal{X}_1$ there.

Similarly, suppose that $a$’s action fully determines the outcome. Then $\mathcal{X}_0 \sim^{(2)}_S \mathcal{X}_3$ so that for all admissible interventions to $(\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2)$ we have

$$r_3(z_0^*, r_1(z_0^*), z_2) - r_3(z_0, r_1(z_0), z_2) = 0,$$
and now we say that $X_1$ and $X_3$ are causally isolated given $X_2$. We can then write $Y_3(\omega) = r_3(\omega_0, Z_1(\omega_1), Z_2(\omega_2)) = \tilde{r}_3(Z_2(\omega_2))$ for some measurable function $\tilde{r}_3$. In particular, the canonical response $Y_2^c$ is given by $Y_2^c = r_3(Y_0^c, Y_1^c, Y_3^c) = \tilde{r}_3(Y_3^c)$, and therefore for any probability measure $P$, $X_0 \not\perp_S X_3$ implies that $Y_1^c \perp Y_3^c$ $| Y_2^c$. A sufficient condition for $X_0 \not\perp_S X_3$ is that $X_0 \not\perp_S X_3$ and $X_1 \not\perp_S X_3$, as depicted in direct causality graph $G_5$.

These examples demonstrate that, for any probability measure $P$, conditional causal isolation is sufficient for conditional independence among canonical responses. It follows that failure of conditional causal isolation is a necessary requirement for conditional dependence for some $P$. In particular, settable variables $X_2$ and $X_3$ must share the principal settable variable $X_0$ as a common cause exclusive of the third variable $X_1$ in order for the canonical responses $Y_2^c$ and $Y_3^c$ to be conditionally dependent given $Y_1^c$. We term this result the conditional Reichenbach principle of common cause.

The unconditional counterpart of this result formally establishes Reichenbach’s principle of common cause in recursive settable systems. This states that, trivially, the principal settable variable $X_0$ must be a common cause of two settable variables say $X_1$ and $X_2$ in order for their canonical responses $Y_1^c$ and $Y_2^c$ to be dependent. Otherwise, either $Y_1^c$ or $Y_2^c$ (or both) is a constant. Observe that we do not need to employ nor do we employ notions of $d$-separation in the above discussion, as these are irrelevant here.

On the other hand, conditional causal isolation is not necessary for conditional independence. The latter may arise for specific probability measures and causal relations. Systems analogous to Markovian and semi-Markovian PCM’s fall into this category (the distinction between Markovian and semi-Markovian systems is that all common causes of canonical responses are observable in the former but not in the latter). Settable systems can accommodate such special systems but need not be confined to them. To illustrate, consider a Markovian-type settable system $S_M$ for the same example as above. In $S_M$, the principal settable variable $X_0$ does not directly cause settable variables $X_1, X_5,$ and $X_6$, corresponding now to $e$’s recommendation, $a$’s action, and the outcome. Instead, $S_M$ assumes the
existence of fundamental settable variables $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$ (so that $\mathcal{S}_M$ has $n=6$ units) such that (see direct causality graph $G_6$):

$$
Y_4(\omega) = r_4(Z_1(\omega)) \\
Y_5(\omega) = r_5(Z_2(\omega), Z_4(\omega)) \\
Y_6(\omega) = r_6(Z_3(\omega), Z_4(\omega), Z_5(\omega_5)).
$$

Finally and importantly, the probability measure $P$ ensures that the canonical settings $Z_1^c$, $Z_2^c$, and $Z_3^c$ are mutually independent (securing the Markov property).

Now consider the probabilistic DAG $G_7$, isomorphic to the subgraph involving $\mathcal{X}_4$, $\mathcal{X}_5$, and $\mathcal{X}_6$ in $G_6$. In $G_7$, we replace settable variables with canonical responses $Y_4^c$, $Y_5^c$, and $Y_6^c$ at the nodes. Arguments similar to those in Section 3.3 demonstrate that conditional causal isolation, together with the independence assumption imposed on $Z_1^c$, $Z_2^c$, and $Z_3^c$ via $P$ imply that the local Markov property holds in DAGs such as $G_7$. Hence, in this particular case, one can apply $d$-separation to learn about conditional independence relations among canonical responses $Y_4^c$, $Y_5^c$, and $Y_6^c$.

For instance, if $\mathcal{X}_4 \overset{D}{\not\rightarrow} \mathcal{S}_M \mathcal{X}_6$ as in direct causality graph $G_8$, we conclude that $Y_4^c \perp Y_6^c \mid Y_5^c$, since $Y_5^c$ $d$-separates $Y_4^c$ and $Y_6^c$ in probabilistic DAG $G_9$. This result formalizes the intuition that conditioning on a variable that fully mediates the effects of a cause on an effect renders the cause and effect conditionally independent. Similarly, if $\mathcal{X}_5 \overset{D}{\not\rightarrow} \mathcal{S}_M \mathcal{X}_6$ as in direct causality graph $G_{10}$, we conclude that $Y_4^c \perp Y_6^c \mid Y_5^c$, since $Y_5^c$ $d$-separates $Y_4^c$ and $Y_6^c$ in probabilistic DAG $G_{11}$, also formalizing the rough intuition that conditioning on a common cause of two variables renders these conditionally independent.

When, as is true here, canonical responses are such that $Y_4^c \perp Y_6^c \mid Y_5^c$ even though $\mathcal{X}_4$ and $\mathcal{X}_6$ are not causally isolated given $\mathcal{X}_5$, we say that $\mathcal{X}_4$ and $\mathcal{X}_6$ are $P$-stochastically isolated given $\mathcal{X}_5$. In such cases, conditional independence holds only for specific probability measures, such as for the Markovian system of the present example.
Next, suppose that $X_4 \neq S_{hit} X_5$ as in direct causality graph $G_{12}$. Then $Y_4^c$ and $Y_5^c$ are not $d$-separated given $Y_6^c$ in probabilistic DAG $G_{13}$. Although lack of $d$-separation does not generally ensure that $Y_4^c \not\perp Y_5^c \mid Y_6^c$, it is often assumed that this is the case. For example, SGS (pp. 35, 56) refer to distributions in which failure of $d$-separation implies conditional dependence as “stable” and Pearl (2000, pp. 48-49;) refers to such distributions as “faithful.” Below, we give general conditions, valid for both Markovian and non-Markovian systems, under which this conclusion holds (see also Wermuth and Cox, 2004). When $Y_4^c \not\perp Y_6^c \mid Y_5^c$ holds, the heuristic intuition that conditioning on a common response induces dependence among independent common causes is made formal.

We emphasize that the assumptions about (1) the existence of fundamental variables $X_1, X_2,$ and $X_3$; (2) the (lack of) causal relations involving these fundamental variables; and (3) the canonical settings $Z_1^c, Z_2^c,$ and $Z_3^c$ being jointly independent are very strong assumptions that need not hold in general and that must be carefully justified if imposed. Nor are these assumptions and graphical criterion assumptions essential to the study of causal and probabilistic relations, as discussed above.

5 Settable Systems

In the next several sections, we formalize and extend the content of the foregoing example to arrive at a general framework connecting causal relations and conditional independence. For
this, we use settable systems, introduced by WC as an extension of the PCM where response functions can arise from optimization, equilibrium, and/or learning. In this section, we briefly describe specialized versions of WC’s definition that are sufficient for our purposes. We refer the interested reader to WC for a detailed discussion of settable systems, their relationship to the PCM (Pearl, 2000), and several examples.

Heuristically, a stochastic settable system is a mathematical framework that describes an environment in which a countable number of units interact under uncertainty. A unit is construed broadly. It could be a neuron, person, machine, firm, market, or a player-decision pair in decision or game theory, for example. There may be a countable infinity of units \( i, i = 1, ..., n \), where \( n \in \mathbb{N}^+ := \mathbb{N} \cup \{\infty\} \) and \( \mathbb{N}^+ \) denotes the positive integers. When \( n = \infty \), we interpret \( i = 1, ..., n \) as \( i = 1, 2, ... \). Random variables are defined on a measurable space \((\Omega, \mathcal{F})\); this provides the foundation for probabilistic statements. For settable systems, it is convenient to define a principal space \( \Omega_0 \) and let \( \Omega := \times_{i=0}^n \Omega_i \), with each \( \Omega_i \) a copy of \( \Omega_0 \). An often convenient choice is \( \Omega_0 = \mathbb{R} \).

In settable systems, there is a settable variable \( X_i \) for each unit \( i \). A settable variable \( X_i \) has a dual aspect. It can be set to a random variable denoted by \( Z_i \) (the setting), where \( Z_i : \Omega_i \rightarrow S_i \) and \( S_i \), the admissible setting values, for \( Z_i \), is a multi-element subset of \( \mathbb{R} \). Or it can be free to respond to settings of other settable variables in the system. In the latter case, it is denoted by the response \( Y_i : \Omega \rightarrow S_i \). The response \( Y_i \) of a settable variable \( X_i \) to the settings of other settable variables is determined by a response function, \( r_i \). For example, \( r_i \) can be determined by a governing principle such as optimization, determining the response for unit \( i \) that is best in some sense, given the settings of other settable variables. The dual role of a settable variable \( X_i : \{0, 1\} \times \Omega \rightarrow S_i \), distinguishing responses \( X_i(0, \omega) := Y_i(\omega) \) and settings \( X_i(1, \omega) := Z_i(\omega) \), \( \omega \in \Omega \), enables us to formalize the directional nature of causal relations, whereby settings of some variables (causes) determine responses of others.

The principal unit \( i = 0 \) plays a key role in understanding and formalizing the connections between probabilistic and causal relations. We let the principal setting \( Z_0 \) and principal response \( Y_0 \) of the principal settable variable \( X_0 \) be such that \( Z_0 : \Omega_0 \rightarrow \Omega_0 \) is the identity map, \( Z_0(\omega_0) := \omega_0 \), and we define \( Y_0(\omega) := Z_0(\omega_0) \). The setting \( Z_0 \) of the principal settable variable may directly influence all other responses in the system, whereas its response \( Y_0 \) is unaffected by other settings. Thus, \( X_0 \) introduces an aspect of “pure randomness” to responses of settable variables.

WC’s definition explicitly accommodates attributes, i.e. fixed objects (e.g., numbers such as \( i \) or sets such as \( S_i \)) associated with each unit \( i \). For conciseness and without essential loss of generality here, we leave attributes implicit.
5.1 Elementary Settable Systems

In elementary settable systems, the response $Y_i$ is determined (actually or potentially) by the settings of all other variables in the system, denoted $Z_{(i)}$. Thus, in elementary settable systems, $Y_i = r_i(Z_{(i)})$. The relation $Y_i = r_i(Z_{(i)})$ corresponds to a structural equation in the classical formulation of systems of structural equations (see, e.g., Heckman, 2005).

The settings $Z_{(i)}$ take values in $S_{(i)} \subseteq \Omega_0 \times \Omega_j \neq \emptyset$. We have $S_{(i)} \subseteq \Omega_0 \times \Omega_j$ if there are joint restrictions on the admissible settings values, such as in “mixed-strategy” static games of complete information, for example, where certain elements of $S_{(i)}$ might represent probabilities that must sum to one (see WC).

We now give a formal definition of elementary settable systems.

**Definition 5.1 Elementary Settable System** Let $(\Omega, \mathcal{F})$ be a measurable space such that $\Omega := \times_{i=0}^{n} \Omega_i$, with each $\Omega_i$ a copy of the principal space $\Omega_0$, containing at least two elements. Let the principal setting $Z_0 : \Omega_0 \rightarrow \Omega_0$ be the identity mapping. For $i = 1, 2, ..., n$, $n \in \mathbb{N}^+$, let $S_i$ be a multi-element Borel-measurable subset of $\mathbb{R}$ and let settings $Z_i : \Omega_i \rightarrow S_i$ be surjective measurable functions. Let $Z_{(i)}$ be the vector including every setting except $Z_i$ and taking values in $S_{(i)} \subseteq \Omega_0 \times \Omega_j \neq \emptyset$. Let response functions $r_i : S_{(i)} \rightarrow S_i$ be measurable functions and define responses $Y_i(\omega) := r_i(Z_{(i)}(\omega))$.

Define settable variables $X_i : \{0, 1\} \times \Omega \rightarrow S_i$ as

$$X_i(0, \omega) := Y_i(\omega) \quad \text{and} \quad X_i(1, \omega) := Z_i(\omega), \quad \omega \in \Omega.$$ 

Define $Y_0$ and $X_0$ by $Y_0(\omega) := X_0(0, \omega) := X_0(1, \omega) := Z_0(\omega), \quad \omega \in \Omega$. Put $\mathcal{X} := \{X_0, X_1, \ldots\}$. The pair $\mathcal{S} := \{\Omega, \mathcal{F}, \mathcal{X}\}$ is an elementary settable system.

A stochastic settable system is thus composed of a “stochastic” component, i.e., the measurable space $(\Omega, \mathcal{F})$, and a structural or causal component $\mathcal{X}$, resting on the stochastic component and consisting of settable variables whose properties are crucially determined by response functions $r := \{r_i\}$.

5.2 Partitioned Settable Systems

In Definition 5.1, a single response $Y_i$ is free to respond to settings of all other variables in the system. We also wish to consider systems in which responses of several settable variables jointly respond to settings of the remaining variables in the system (see e.g. Wermuth and Cox, 2004). This can occur, for example, when responses are determined as a solution to a joint optimization problem. Such specifications are formally implemented in settable
systems by partitioning the system under study to group jointly responding variables into specific blocks. The system in Definition 5.1 is called “elementary,” as every unit \( i \) forms a block by itself. We now define general partitioned settable systems.

**Definition 5.2 Partitioned Settable System** Let \((\Omega, \mathcal{F}), X_0, n, \) and \( S_i, i = 1, \ldots, n \), be as in Definition 5.1. Let \( \Pi = \{\Pi_b\} \) be a partition of \( \{1, \ldots, n\} \), with cardinality \( B \in \mathbb{N}^+ \) \((B := \#\Pi)\). For \( i = 1, 2, \ldots, n \), let \( Z_i^\Pi \) be settings and let \( Z_i^\Pi \) be the vector containing \( Z_0^\Pi \) and \( Z_i^\Pi \), \( i \notin \Pi_b \), and taking values in \( S_i^\Pi \subseteq \Omega_0 \times \prod_{i \in \Pi_b} S_i \), \( S_i^\Pi \neq \varnothing \), \( b = 1, \ldots, B \). For \( b = 1, \ldots, B \) and \( i \in \Pi_b \), suppose there exist measurable functions \( r_i^\Pi: S_i^\Pi \to S_i \), specific to \( \Pi \) such that responses \( Y_i^\Pi(\omega) \) are jointly determined as

\[
Y_i^\Pi := r_i^\Pi(Z_i^\Pi).
\]

Define the settable variables \( X_i^\Pi : \{0, 1\} \times \Omega \to S_i \) as

\[
X_i^\Pi(0, \omega) := Y_i^\Pi(\omega) \quad \text{and} \quad X_i^\Pi(1, \omega) := Z_i^\Pi(\omega), \quad \omega \in \Omega.
\]

Put \( \mathcal{X}^\Pi := \{X_0^\Pi, X_1^\Pi, X_2^\Pi \ldots\} \). The pair \( (\Omega, \mathcal{F}), (\Pi, \mathcal{X}^\Pi) \) is a partitioned settable system.

The settings \( Z_i^\Pi \) are allowed to be partition-specific; this is especially relevant when the admissible set \( S_i^\Pi \) imposes restrictions on the admissible values of \( Z_i^\Pi \). Crucially, response functions and responses are partition-specific. In Definition 5.2, the joint response function \( r_i^\Pi := (r_i^\Pi, i \in \Pi_b) \) specifies how the settings \( Z_i^\Pi \) outside of block \( \Pi_b \) determine the joint response \( Y_i^\Pi := (Y_i^\Pi, i \in \Pi_b) \), i.e., \( Y_i^\Pi = r_i^\Pi(Z_i^\Pi) \).

Below, it will also be convenient to let \( \Pi_0 = \{0\} \) represent the block corresponding to the principal settable variable.

### 5.3 Recursive Settable Systems

In what follows, we often consider recursive partitioned settable systems, defined next. For \( 0 \leq a \leq b \), we define \( \Pi_{[a:b]} := \Pi_a \cup \ldots \cup \Pi_{b-1} \cup \Pi_b \). (For \( a < b \), \( \Pi_{[b:a]} := \varnothing \).

**Definition 5.3 Recursive Partitioned Settable System** Let \( S \) be a partitioned settable system. For \( b = 0, 1, \ldots, B \), let \( Z_i^\Pi_{[b:b]} \) denote the vector containing the settings \( Z_i^\Pi \) for \( i \in \Pi_{[0:b]} \) and taking values in \( S_i^\Pi_{[b:b]} \subseteq \Omega_0 \times \prod_{i \in \Pi_{[1:b]}} S_i \), \( S_i^\Pi_{[0:b]} \neq \varnothing \). For \( b = 1, \ldots, B \) and \( i \in \Pi_b \), suppose that \( r_i^\Pi := \{r_i^\Pi\} \) is such that the responses \( Y_i^\Pi = X_i^\Pi(1, \cdot) \) are determined as

\[
Y_i^\Pi := r_i^\Pi(Z_i^\Pi_{[0:b-1]}).
\]
Then we say that \( \Pi \) is a **recursive partition**, \( r^\Pi \) is recursive, and the pair \( S := \{ (\Omega, \mathcal{F}), (\Pi, \mathcal{X}^\Pi) \} \) is a **recursive partitioned settable system** or simply that \( S \) is recursive.

We employ the convenient structure of recursive systems to provide definitions of indirect and total causality. This also facilitates the comparison between our results and the DAG-related literature. We leave the study of the interrelations between (conditional) independence and indirect and total causal relationships in non-recursive systems (see, e.g., Lauritzen and Richardson, 2002) for other work.

# 6 Causality in Settable Systems

Settable systems provide a suitable framework for the study of causality. We now give definitions of several notions of causality within this framework, based on functional dependence: direct causality, indirect causality via and exclusive of a given set of variables, and total causality. These notions refine and extend related concepts referenced below. Of particular note is that we define causality in terms of settable variables rather than random variables or events, as is typical elsewhere. For notational convenience, we may suppress explicit reference in what follows to the superscript \( \Pi \) in \( Z_i^\Pi, r_i^\Pi, Y_i^\Pi, \) and \( \mathcal{X}_i^\Pi; \) it should nevertheless be borne in mind that these functions are partition-specific.

## 6.1 Direct Causality and Direct Causality Graphs

Direct causality can be defined for both recursive and non-recursive settable systems. Heuristically, we say that a settable variable \( \mathcal{X}_i, i \notin \Pi_b \), directly causes \( \mathcal{X}_j, j \in \Pi_b \), in \( S \) when the response for \( \mathcal{X}_j \) differs for different settings in \( \mathcal{X}_i \), while holding all other variables corresponding to units outside of \( \Pi_b \) to the same setting values. There are two main ingredients to this notion of direct causality. Let \( z_{(b)(i)} \) denote the vector containing all elements of setting values \( z_{(b)} \) except \( z_i \). The first ingredient is an *admissible intervention*, \( (z_{(b)(i)}; z_i) \rightarrow (z_{(b)(i)}; z_i^*) \). We define this to be a pair of elements of \( S_{(b)} \), i.e., \( (z_{(b)(i)}; z_i) \rightarrow (z_{(b)(i)}; z_i^*) := ((z_{(b)(i)}; z_i), (z_{(b)(i)}; z_i^*)) \), where we abuse notation somewhat by reordering the vector arguments for convenience. The intervention references only setting values corresponding to units outside of \( \Pi_b \). Note also that it differs only in the final component. The second ingredient is the behavior of the response to this intervention.

We formalize this notion of direct causality as follows.

**Definition 6.1 Direct Causality** Let \( S \) be a partitioned settable system. For given positive integer \( b \), let \( j \in \Pi_b \). (i) For given \( i \notin \Pi_b \), \( \mathcal{X}_i \) **directly causes** \( \mathcal{X}_j \) in \( S \) if there exists

\[ \cdots \]
an admissible intervention \((z_{(b)(i)}, z_i) \rightarrow (z_{(b)(i)}, z^*_i)\) such that

\[ r_j(z_{(b)(i)}, z^*_i) - r_j(z_{(b)(i)}, z_i) \neq 0, \]

and we write \(X_i \overset{D}{\leftrightarrow}_S X_j\). Otherwise, we say \(X_i\) does not directly cause \(X_j\) in \(S\) and write \(X_i \not\overset{D}{\leftrightarrow}_S X_j\). (ii) For \(i, j \in \Pi_b\), \(X_i \not\overset{D}{\leftrightarrow}_S X_j\).

We emphasize that even though we follow the literature in referring to “interventions,” with their mechanistic or manipulative connotations, the mathematical concept only involves the properties of a response function on its domain.

According to this definition, direct causality may fail either because the set \(S^{(b)}_\Pi\) is so constrained that it does not possess an admissible intervention of the desired form, or because it does, but the response is the same for both elements of every admissible intervention of the specified form. The latter is perhaps the more common or intuitively appealing possibility, but we need not distinguish further between these possibilities.

Note that, by definition, variables within the same block do not directly cause each other. In particular \(X_i \not\overset{D}{\leftrightarrow}_S X_i\). Also, Definition 6.1 permits mutual causality, so that \(X_i \overset{D}{\leftrightarrow}_S X_j\) and \(X_j \overset{D}{\leftrightarrow}_S X_i\) without contradiction for \(i\) and \(j\) in different blocks. Mutual causality is ruled out in SGS (p. 42), for example, where it is an axiom that if \(A\) causes \(B\) then \(B\) does not cause \(A\).

An important aspect of Definition 6.1 is the explicit causal role permitted for \(X_0\). The background variables of the PCM are analogous to \(X_0\), as they are not determined by other system variables, but background variables explicitly cannot act as causes in the PCM.

We call the response value difference in Definition 6.1 the direct effect of \(X_i\) on \(X_j\) in \(S\) of the specified intervention. This corresponds to the notion of “controlled” direct effect in Pearl (2001). Nevertheless, the PCM requires a unique fixed point, a requirement absent here; the PCM also does not have a notion of partitioning, so the PCM notion pertains only to elementary partitions; and the PCM does not account for possible joint restrictions on setting values, and thus effectively assumes that \(S^{(b)}_\Pi = \Omega_0 \times_{i \neq j} S_i\).

Direct causality relations have a convenient graphical representation. For this, we introduce notions of paths, successors, predecessors, and intercessors, adapting graph theoretic concepts discussed, for example, by Bang-Jensen and Gutin (2001).

**Definition 6.2 Paths, Successors, Predecessors, and Intercessors** Let \(S\) be a partitioned settable system. For given positive integer \(b\) let \(j \in \Pi_b\) and \(i \notin \Pi_b\). We call the collection of settable variables \(\{X_i, X_{i_1}, \ldots, X_{i_m}, X_j\}\) an \((X_i, X_j)\)-\textbf{walk} of length \(m + 1\) if
When the elements of an \((X_i, X_j)\)-walk are distinct, we call it an \((X_i, X_j)\)-path. We say \(X_i\) precedes \(X_j\) or \(X_j\) succeeds \(X_i\) if there exists at least one \((X_i, X_j)\)-path of positive length. If \(X_i\) precedes \(X_j\), we call \(X_i\) a predecessor of \(X_j\), and we call \(X_j\) a successor of \(X_i\). If \(X_i\) precedes \(X_j\) and \(X_i\) succeeds \(X_j\), we say \(X_i\) and \(X_j\) belong to a cycle. If \(X_i\) and \(X_j\) do not belong to a cycle, \(X_k\) succeeds \(X_i\), and \(X_k\) precedes \(X_j\), we say \(X_k\) intercedes \(X_i\) and \(X_j\). If \(X_k\) intercedes \(X_i\) and \(X_j\), we call \(X_k\) an \((X_i, X_j)\)-intercessor. We denote by \(I_{i,j}\) the set of \((X_i, X_j)\)-intercessors.

The direct causality graph for a given partitioned settable system \(S\) is a directed graph \(G := (V, E)\) with a non-empty countable set of vertices \(V = \{X_i : i = 0, 1, \ldots, n\}\) and a set of arcs \(E \subseteq V \times V\) of ordered pairs of distinct vertices such that an arc \((X_i, X_j)\) belongs to \(E\) if and only if \(X_i \xrightarrow{D} S X_j\). From Definition 6.1, there exists at most one \((X_i, X_j)\) arc, so \(G\) need not contain nor can it contain “parallel arcs.” Since \(X_i \xrightarrow{D} S X_i\), there can be no arc \((X_i, X_i)\) in \(E\), so \(G\) need not and can not contain self-directed arcs or “loops.”\(^5\)

Direct causality graph \(G_{14}\) illustrates the concepts of Definition 6.2. We have that \(\{X_0, X_1, X_2, X_3, X_1, X_2, X_3, X_4\}\) and \(\{X_0, X_1, X_2, X_3, X_4\}\) are an \((X_0, X_4)\)-walk of length 7 and an \((X_0, X_4)\)-path of length 4, respectively. We also have that \(X_0\) precedes \(X_4\), \(X_3\) succeeds \(X_1\), and that \(X_1\) and \(X_3\) belong to a cycle, as do \(X_4\) and \(X_5\). The set of \((X_1, X_4)\)-intercessors is given by \(I_{1:4} = \{X_2, X_3\}\). We use the term “intercessor” instead of the possible descriptor “mediator,” as the latter may connote transmission; we want to avoid this, because intercessors need not transmit effects, as we further explain below.

\[
\begin{array}{c}
X_0 \\
\xrightarrow{X_1} \\
\xrightarrow{X_3} \\
\xrightarrow{X_4} \\
\xrightarrow{X_5} \\
\end{array}
\]

\(\text{Graph 14 (G}_{14}\)

We emphasize that these direct causality graphs differ from other graphs in the literature. Nodes in direct causality graphs represent settable variables, not random variables or events; arcs represent direct causality relations, not functional or probabilistic dependence.

\(^5\)Loops and parallel arcs can nevertheless be useful in other contexts; see, for example, Golubitsky and Stewart (2006). With loops or parallel arcs permitted, one may have a “directed pseudograph” or a “directed multigraph” (see Bang-Jensen and Gutin, 2001, p. 4). These are not relevant here.
6.1.1 Direct Causality in Recursive Settable Systems

We now consider how Definition 6.1 (direct causality) specializes to recursive systems. For this, let $0 \leq b_1 < b_2$ and take $i \in \Pi_{b_1}$ and $j \in \Pi_{b_2}$. We write values of settings corresponding to $\Pi_{[a:b]}$ as $z_{[a:b]}$. We also let $z_{[0:b]}(i)$ denote a vector of values for settings for all settable variables corresponding to $\Pi_{[0:b]}$ except $\mathcal{X}_i$. Since $\mathcal{S}$ is recursive, we can express response values for $\mathcal{X}_j$ as $r_j(z_{[0:b_2-1]})$. We abuse notation somewhat to permute the arguments of $r_j$ in a way that emphasizes their recursive relation to the argument corresponding to $\mathcal{X}_i$. In particular, we write

$$r_j\left(z_{[0:b_1]}(i), z_i, z_{[b_1+1:b_2-1]}\right) = r_j\left(z_{[0:b_2-1]}\right).$$

Definition 6.1 then concludes that $\mathcal{X}_i \xrightarrow{D} \mathcal{X}_j$ if there exists an admissible intervention \( \left(z_{[0:b_1]}(i), z_i, z_{[b_1+1:b_2-1]}\right) \rightarrow \left(z_{[0:b_1]}(i), z_i^*, z_{[b_1+1:b_2-1]}\right) \) such that

$$r_j\left(z_{[0:b_1]}(i), z_i^*, z_{[b_1+1:b_2-1]}\right) - r_j\left(z_{[0:b_1]}(i), z_i, z_{[b_1+1:b_2-1]}\right) \neq 0.$$

Clearly if $\mathcal{S}$ is recursive, successors do not directly cause predecessors; that is, if $i \in \Pi_{b_1}$ and $j \in \Pi_{b_2}$ with $b_2 < b_1$, then $\mathcal{X}_i \nRightarrow \mathcal{X}_j$. In particular, if $\mathcal{X}_i \xrightarrow{D} \mathcal{X}_j$ then $\mathcal{X}_j \xrightarrow{D} \mathcal{X}_i$. Thus, recursive systems do not admit mutual causality. For the direct causality graph, this means that we cannot have both arcs $(\mathcal{X}_i, \mathcal{X}_j)$ and $(\mathcal{X}_j, \mathcal{X}_i)$ belonging to $E$. In addition, a recursive system $\mathcal{S}$ is acyclic: it does not admit cycles of the form $\mathcal{X}_i \xrightarrow{D} \mathcal{X}_{i_1} \xrightarrow{D} \mathcal{X}_{i_2} \ldots \xrightarrow{D} \mathcal{X}_{i_m} \xrightarrow{D} \mathcal{X}_i$. Thus, when $\mathcal{S}$ is recursive, its corresponding direct causality graph $G$ is a DAG.

In the expression above for recursive settable system direct causality, the values for successors to $\mathcal{X}_i$ (corresponding to blocks $\Pi_{[b_1+1:b_2-1]}$) are set to the same arbitrary value $z_{[b_1+1:b_2-1]}$ in both argument lists. Sometimes it is of interest to evaluate the direct effect on $\mathcal{X}_j$ of $\mathcal{X}_i$ when values of $\mathcal{X}_i$’s successors are set in both argument lists to the response value obtained when $\mathcal{X}_i$ is set to $z_i$. When a setting is given by the response to its predecessors’ settings, we call it canonical. Thus, the canonical setting for $\mathcal{X}_i$, $i \in \Pi_b$, is

$$Z_i^c = Y_i := r_i(Z_{[0:b-1]}).$$

Then setting values $Z_i^c_{[b_1+1:b_2-1]}$ determined as responses $y_{[b_1+1:b_2-1]}$ to the admissible values of their predecessors’ settings are

$$y_{[b_1+1:b_2-1]} = r_{[b_1+1:b_2-1]}(z_{[0:b_1]}).$$

The elements of this response vector obtain by recursive substitution. Any given element of this vector depends only on its corresponding predecessors. The direct effect associated with this configuration is then evaluated as

$$r_j\left(z_{[0:b_1]}(i), z_i^*, y_{[b_1+1:b_2-1]}\right) - r_j\left(z_{[0:b_1]}(i), z_i, y_{[b_1+1:b_2-1]}\right).$$
Pearl (2001) refers to this as the “natural” direct effect. Although Pearl (2001) does not assume recursiveness, he employs the PCM, with its unique fixed point requirement. As mentioned above, we do not require a fixed point, unique or otherwise, so just as for our prior notion of direct causality, this concept of direct causality does not depend on this.

### 6.1.2 Relation to Other Notions of Direct Causality

Now consider the following system of three settable variables (see direct causality graph $G_{15}$) to illustrate the relationships between our Definition 6.1 of direct causality and several other notions of direct effects discussed in the literature.

\[
\mathcal{X}_1(0, \cdot) = r_1(\mathcal{X}_0(1, \cdot))
\]

\[
\mathcal{X}_2(0, \cdot) = r_2(\mathcal{X}_0(1, \cdot), \mathcal{X}_1(1, \cdot)).
\]

Definition 6.1 concludes that $\mathcal{X}_0 \xrightarrow{D} \mathcal{X}_2$ if there exists an admissible intervention $(z_0, z_1) \rightarrow (z_0^*, z_1)$ such that

\[
r_2(z_0^*, z_1) - r_2(z_0, z_1) \neq 0.
\]

When this difference is non-zero, it justifies the link from $\mathcal{X}_0$ to $\mathcal{X}_2$. This difference corresponds to the notion of “controlled direct effect” in Pearl (2001).

If $z_1$ is restricted to a specific suitable value, then we obtain a notion in the spirit of the “standardized direct effect” of Didelez, Dawid, and Geneletti (2006) and Geneletti (2007). In particular, the canonical choice $z_1^* = r_1(z_0)$ yields Pearl’s (2001) previously mentioned natural direct effect

\[
r_2(z_0^*, r_1(z_0)) - r_2(z_0, r_1(z_0)).
\]

This also is what Robins and Greenland (1992) and Robins (2003) call the “pure” direct effect. These same authors refer to

\[
r_2(z_0^*, r_1(z_0^*)) - r_2(z_0, r_1(z_0^*))
\]

as the “total direct effect.”

In other cases, the literature considers notions of direct effects defined as a contrast in some aspect of the distributions of responses for different settings. For example, let $P$ be
a probability measure on \((\Omega, \mathcal{F})\); then the “average” direct effect of \(X_1\) on \(X_2\) in the above example is given by

\[
E[r_2(Z_0, z_1^*) - r_2(Z_0, z_1)],
\]

where \(E\) is the expectation operator associated with \(P\).

Here, we consider direct effects to be differences in response values for any admissible intervention of the specified form. As Holland (1986) notes, these effects need not be identifiable absent other assumptions. Nevertheless, the direct causality concept of Definition 6.1 is in a precise sense the simplest and most general of the alternatives discussed. It is simplest, in that direct causality is well defined even in the absence of recursive structure or fixed points. It is most general, as it is necessary but not sufficient for the others.

### 6.2 Indirect Causality in Settable Systems

We next define notions of indirect causality for recursive systems. We distinguish notions of indirect causality via and exclusive of specified variables. These definitions extend notions of indirect causality in Robins and Greenland (1992), SGS, Pearl (2001), Robins (2003), Didelez, Dawid, and Geneletti (2006), and Geneletti (2007), and notions of “path-specific” effects in Pearl (2001) and Avin, Shpitser, and Pearl (2005). Although these extensions are of interest in their own right, their greater significance is that they provide appropriate tools for establishing the conditional Reichenbach principle of common cause, as well as later results on \(d\)–separation and \(D\)–separation.

#### 6.2.1 Indirect Causality Via Given Variables

**Motivating Examples** The basic idea of indirect causality adopted here is straightforward. Consider, for example, the system illustrated in \(G_{15}\). There, \(X_0\) indirectly causes \(X_2\) via \(X_1\) if there exists an admissible intervention \((z_0, r_1(z_0)) \rightarrow (z_0, r_1(z_0^*))\) such that

\[
r_2(z_0, r_1(z_0^*)) - r_2(z_0, r_1(z_0)) \neq 0.
\]

In the first case, \(Z_0\) is set to the value \(z_0\) and \(Z_1\) to the canonical value \(r_1(z_0)\). In the second case, \(Z_0\) is set to the value \(z_0\) and \(Z_1\) is set to the canonical value \(r_1(z_0^*)\) that obtains when \(Z_0\) is set to \(z_0^*\). This corresponds to the notion of “natural indirect effect” in Pearl (2001) and Didelez, Dawid, and Geneletti (2006) and to the notion of “pure indirect effect” in Robins and Greenland (1992) and Robins (2003).

It is necessary but not sufficient for our notion of indirect causality that \(X_0\) directly cause \(X_1\) and that \(X_1\) directly cause \(X_2\). We emphasize that transitivity of causation is
not guaranteed here, unlike classical treatments such as SGS (p. 42), where transitivity of causation is axiomatic. Instead, transitivity depends on the response functions. For example, if \( r_1(z_0) = \max(z_0, 0) \) and \( r_2(z_0, z_1) = \min(z_1, 0) \), then \( X_0 \xrightarrow{D} S X_1 \) and \( X_1 \xrightarrow{D} S X_2 \), but \( X_0 \) does not indirectly cause \( X_2 \), as \( r_2(z_0, r_1(z_0^*)) = \min(\max(z_0^*, 0), 0) = 0 \) for all \( z_0^* \).

With transitivity, \( X_i \) is an indirect cause of \( X_j \) if there exists an \( (X_i, X_j) \)-path of length greater than 2 (SGS, pp. 44-45). Although this example conveys the basic idea, we work with more refined notions of indirect causality, elaborated below.

In \( G_{15} \), \( X_1 \) is the only \( (X_0, X_2) \)-intercessor. In the presence of multiple intercessors, we may be interested in indirect causality via just one specified variable. Consider, for example, the system illustrated in \( G_{16} \).

![Graph 16](image)

We say that \( X_0 \) indirectly causes \( X_3 \) via \( X_1 \) if there exists an admissible intervention \( (z_0, r_1(z_0), z_2) \rightarrow (z_0, r_1(z_0^*), z_2) \) such that

\[
r_3(z_0, r_1(z_0^*), z_2) - r_3(z_0, r_1(z_0), z_2) \neq 0.
\]

If we restrict \( z_2 \) to the value \( r_2(z_0) \) in the above difference, we essentially obtain the “path-specific effect transmitted through the path \{X_0, X_1, X_3\}” in Pearl (2001) and Avin, Shpitser, and Pearl (2005).

More generally, we may consider notions of indirect causality via not just one but several settable variables, as illustrated in \( G_{17} \).

![Graph 17](image)

Here, we say that \( X_0 \) indirectly causes \( X_4 \) via \( X_1 \) or \( X_3 \) if there exist an admissible intervention \( (r_2(z_0, r_1(z_0)), r_3[r_1(z_0), r_2\{z_0, r_1(z_0)\}], r_4[r_2(z_0^*, r_1(z_0^*)), r_3[r_1(z_0^*), r_2\{z_0^*, r_1(z_0^*)\}]) \) such
that

\[ r_4(r_2(z_0, r_1(z_0^*)), r_3(r_1(z_0^*), r_2(z_0^*))) - r_4(r_2(z_0, r_1(z_0)), r_3[r_1(z_0), r_2(z_0, r_1(z_0))]) \neq 0. \]

(Note that here and elsewhere we simplify notation by omitting response function arguments corresponding to variables that are not direct causes of the specified response.) Setting the first arguments of \( r_4 \) to \( r_2(z_0, r_1(z_0^*)) \) and \( r_2(z_0, r_1(z_0)) \) in the response functions above ensures that the difference in the response for \( \mathcal{X}_4 \) is not due to effects transmitted through the path \( \{\mathcal{X}_0, \mathcal{X}_2, \mathcal{X}_4\} \).

**The General Case** In the general case, the idea underlying indirect causality in recursive systems is essentially the same as in these examples, but to express this rigorously demands careful attention to a perhaps daunting mass of detail. Roughly speaking, however, we say that \( \mathcal{X}_i \) indirectly causes \( \mathcal{X}_j \) via \( (\mathcal{X}_i, \mathcal{X}_j) \)—intercessors \( \mathcal{X}_A \), if the response of \( \mathcal{X}_j \) differs when the effects of setting \( \mathcal{X}_i \) to the value \( z_i \) as opposed to \( z_i^* \) are not transmitted directly, but only through \( \mathcal{X}_A \).

In order to study the response of \( \mathcal{X}_j \) under the relevant scenarios, we partition the \( (\mathcal{X}_i, \mathcal{X}_j) \)–intercessors in a recursive manner relative to \( \mathcal{X}_A \). We distinguish the \( (\mathcal{X}_i, \mathcal{X}_j) \)–intercessors that belong to paths through \( \mathcal{X}_A \) from those that don’t. Among the former, we distinguish: (i) the variables that strictly precede \( \mathcal{X}_A \); (ii) \( \mathcal{X}_A \); (iii) the variables that intercede elements of \( \mathcal{X}_A \); and (iv) the variables that strictly succeed \( \mathcal{X}_A \).

For illustration, we employ system \( \mathcal{S}_{18} \), with direct causality relations illustrated in graph \( G_{18} \), where \( \Pi_1 = \{1, 2\} \) and \( \Pi_b = \{b + 1\} \) for \( b = 2, ..., 8 \). The complexity of this example is not capricious. This is the simplest system permitting a full illustration of the relationships that must be considered in a general definition of indirect causality.

![Graph 18 (G_{18})](image-url)
To begin the illustration, take $b_1 < b_2$, $i \in \Pi_{b_1}$, $j \in \Pi_{b_2}$. For example, in $S_{18}$, let $b_1 = 1$ and $b_2 = 8$, let $i = 2$ (the second element of $\Pi_1 = \{1, 2\}$), and let $j = 9$ (the sole element of $\Pi_8$). We denote by $\text{ind}(I_{i,j})$ the indexes of the elements of the $(X_i, X_j)$—intercessors $I_{i,j}$. For example, in $S_{18}$, we have $\text{ind}(I_{2,9}) = \{3, 4, 5, 6, 7, 8\}$. We treat elements of $\Pi_{[0:b_2]}$ that do not correspond to $(X_i, X_j)$—intercessors as elements of $\Pi_{[0:b_1]}$ or $\Pi_{b_2}$ without loss of generality. Here, $\text{ind}(I_{i,j}) = \Pi_{[b_1+1:b_2-1]}$.

Let $A$ be a subset of $\text{ind}(I_{i,j})$. In $S_{18}$, we can let $A = \{5, 7\}$, say. In what follows we order the arguments of response values $r_j(z_{[0:b_2-1]} | x_j)$ for $x_j$ to emphasize their recursive ordering in relation to $X_i$ and $X_A$.

For given $k \in A$, let $I_{i,j}^k := I_{i:k} \cup \{X_k\} \cup I_{k:j}$ denote the $(X_i, X_j)$—intercessors for paths through $X_k$, and for $X_A := \cup_{k \in A} \{X_k\}$, let $I_{i,j}^A := \cup_{k \in A} I_{i,j}^k$ denote the $(X_i, X_j)$—intercessors for paths through $X_A$. (For $A = \emptyset$ we let $I_{i,j}^A = \emptyset$.) Thus, in $S_{18}$ we have $\text{ind}(I_{i,j}^5) = \text{ind}(I_{2,9}) = \{3, 5, 6, 7, 8\}$ and it follows that $\text{ind}(I_{i,j}^A) = \{3, 5, 6, 7, 8\}$ as well.

Let $X_{\overline{A}} := I_{i,j} \setminus I_{i,j}^A$ denote the $(X_i, X_j)$—intercessors not belonging to paths through $X_A$, and let $\overline{A}$ denote the set of indexes of the elements of $X_{\overline{A}}$. In system $S_{18}$, $A = \text{ind}(I_{i,j}) \setminus \text{ind}(I_{i,j}^A) = \{3, 4, 5, 6, 7, 8\} \setminus \{3, 5, 6, 7, 8\} = \{4\}$. Thus, we have $\text{ind}(I_{i,j}) = \text{ind}(I_{i,j}^A) \cup A$ and $\text{ind}(I_{i,j}^A) \cap A = \emptyset$.

We now partition $\text{ind}(I_{i,j}^A)$ into four mutually exclusive and collectively exhaustive subsets. First, let $X_{\overline{A}} := \cup_{k \in A} I_{k:i} \setminus X_A$ denote the inter-$X_A$ intercessors excluded from $X_A$, and let $\overline{A}$ denote the set of indexes of the elements of $X_{\overline{A}}$. In $S_{18}$, we have $\overline{A} = \{6\}$.

Next, we distinguish between the $(X_i, X_j)$—intercessors for paths through $X_A$ that strictly precede or succeed $X_A$. We define the $X_A$ predecessors excluded from $X_A \cup X_{\overline{A}}$:

\[ P_{i,j}^A := \cup_{k \in A} \{X_i \in I_{i,j}^A \text{ and } X_i \notin (X_A \cup X_{\overline{A}}) : X_i \text{ precedes } X_k\}, \]

and the $X_A$ successors excluded from $X_A \cup X_{\overline{A}}$:

\[ S_{i,j}^A := \cup_{k \in A} \{X_i \in I_{i,j}^A \text{ and } X_i \notin (X_A \cup X_{\overline{A}}) : X_i \text{ succeeds } X_k\}. \]

In the example illustrated in $G_{18}$, we have $\text{ind}(P_{i,j}^A) = \{3\}$ and $\text{ind}(S_{i,j}^A) = \{8\}$.

By construction, $\text{ind}(I_{i,j}^A) = \text{ind}(P_{i,j}^A) \cup A \cup \overline{A} \cup \text{ind}(S_{i,j}^A)$, and these subsets are mutually exclusive. In our example, $\text{ind}(I_{i,j}^A) = \{3, 5, 6, 7, 8\}$ and $\text{ind}(P_{i,j}^A) \cup A \cup \overline{A} \cup \text{ind}(S_{i,j}^A) = \{3\} \cup \{5, 7\} \cup \{6\} \cup \{8\}$. Thus, $\text{ind}(P_{i,j}^A), A, \overline{A}$, and $\text{ind}(S_{i,j}^A)$ partition $\text{ind}(I_{i,j})$.

We now use this partition to represent response values for $X_j$ in a convenient form. Recall that $z_{[0:b_1]}(i)$ denotes a vector of values for settings for the vector of settable variables $X_{[0:b_1]}(i)$ corresponding to $\Pi_{[0:b_1]} \setminus \{i\}$. Thus, in $S_{18}$, $z_{[0:1]}(2)$ denotes values of settings for $X_0$ and $X_1$. Similarly, let $z_{i:A}$, $z_{\overline{A}}$, $z_A$, $z_{\overline{A}}$, and $z_{A:j}$ denote vectors of values of settings for
elements of $\mathcal{P}_{i,j}^A$, $\mathcal{X}_A$, $\mathcal{X}_A$, and $\mathcal{S}_{i,j}$ respectively. We now slightly abuse notation to
represent response values for $\mathcal{X}_j$ (recall $j \in \Pi_{b_2}$) as

$$r_j(z_{[0:b_1](i)}, z_i, z_{i:A}, z_A, z_{\overline{A}}, z_{A:j}) = r_j(z_{[0:b_2-1]}),$$

where the arguments of $r_j$ have been reordered in a particular way, so as to focus attention
on settings $z_i$ and $z_A$ of $\mathcal{X}_j$ and $\mathcal{X}_A$.

Observe that when $A = \text{ind} (\mathcal{I}_{i,j})$, the sets $\text{ind} (\mathcal{P}_{i,j}^A)$, $\overline{A}$, and $\text{ind} (\mathcal{S}_{i,j}^A)$ are empty, and
we write $r_j(z_{[0:b_1](i)}, z_i, z_A) = r_j(z_{[0:b_2-1]}).$ Alternatively, when $A = \emptyset$, the sets $\text{ind} (\mathcal{P}_{i,j}^A)$,
$\overline{A}$, and $\text{ind} (\mathcal{S}_{i,j}^A)$ are empty, whereas $A = \text{ind} (\mathcal{I}_{i,j})$, and we write $r_j(z_{[0:b_1](i)}, z_i, z_A) =

r_j(z_{[0:b_2-1]}).

We use the recursiveness of $\mathcal{S}$ and the definitions above to represent vectors of response
values for elements of $\mathcal{P}_{i,j}^A$, $\mathcal{X}_A$, $\mathcal{X}_A$, and $\mathcal{S}_{i,j}$ respectively in the following form, useful
for general definitions of indirect causality:

$$r_{i:A}(z_{[0:b_1](i)}, z_i)$$
$$r_{\overline{A}}(z_{[0:b_1](i)}, z_i, z_{i:A})$$
$$r_A(z_{[0:b_1](i)}, z_i, z_{i:A}, z_{\overline{A}})$$
$$r_{\overline{A}}(z_{[0:b_1](i)}, z_i, z_{i:A}, z_A)$$

and

$$r_{A:j}(z_{[0:b_1](i)}, z_i, z_{i:A}, z_A, z_{A:j}).$$

Here too, the elements of these response vectors obtain by recursive substitution. Any
given element of one of these vectors depends only on its predecessors. Thus, although $z_A$
appears as an argument in $r_{\overline{A}}$, only the predecessor elements of $z_A$ for a given response
determine that response. By definition, an element of $\mathcal{X}_A$ can not directly cause elements
of $\mathcal{P}_{i,j}^A$, $\mathcal{X}_A$, or $\mathcal{X}_A$, nor can it be directly caused by elements of $\mathcal{X}_A$, $\mathcal{X}_A$, or $\mathcal{S}_{i,j}^A$.

Finally, we introduce a notation for canonical settings defined as responses to specific
setting values:

$$y_{i:A} = r_{i:A}(z_{[0:b_1](i)}, z_i)$$
$$y_{\overline{A}} = r_{\overline{A}}(z_{[0:b_1](i)}, z_i, y_{i:A})$$
$$y_A = r_A(z_{[0:b_1](i)}, z_i, y_{i:A}, y_{\overline{A}})$$
$$y_{\overline{A}} = r_{\overline{A}}(z_{[0:b_1](i)}, z_i, y_{i:A}, y_A)$$

$$y_{i:A}^* = r_{i:A}(z_{[0:b_1](i)}, z_i^*)$$
$$y_{\overline{A}}^* = r_{\overline{A}}(z_{[0:b_1](i)}, z_i^*, y_{i:A}^*)$$
$$y_A^* = r_A(z_{[0:b_1](i)}, z_i^*, y_{i:A}^*, y_{\overline{A}}^*)$$
$$y_{\overline{A}}^* = r_{\overline{A}}(z_{[0:b_1](i)}, z_i^*, y_{i:A}^*, y_A^*)$$

We can now state our first formal definition of indirect causality.
Definition 6.3 Indirect Causality via $\mathcal{X}_A$ Let $S$ be recursive. For given non-negative integers $b_1$ and $b_2$ with $b_1 < b_2$, let $i \in \Pi_{b_1}$, let $j \in \Pi_{b_2}$, and let $A$ be a subset of $\text{ind}(\mathcal{I}_{ij})$. Then $\mathcal{X}_i$ indirectly causes $\mathcal{X}_j$ via $\mathcal{X}_A$ in $S$ if there exists an admissible intervention to $(\mathcal{X}_{[0:b_1](i)}, \mathcal{X}_i, \mathcal{P}_{ij}^A, \mathcal{X}^A, \mathcal{X}_{j}^A, S_{ij}^A)$ with corresponding responses for $\mathcal{X}_j$ such that

$$r_j(z_{[0:b_1](i)}, z_i, z_i: A, z_A, y_A^*, r_A(z_{[0:b_1](i)}, z_i, z_i: A, y_A^*)),$$

$$r_A(z_{[0:b_1](i)}, z_i, z_i: A, z_A, y_A^*, r_A(z_{[0:b_1](i)}, z_i, z_i: A, y_A^*))$$

and we write $\mathcal{X}_i \overset{I[A]}{\Rightarrow}_S \mathcal{X}_j$. Otherwise, we say that $\mathcal{X}_i$ does not indirectly cause $\mathcal{X}_j$ via $\mathcal{X}_A$ in $S$ and we write $\mathcal{X}_i \not\overset{I[A]}{\Rightarrow}_S \mathcal{X}_j$. When $A = \text{ind}(\mathcal{I}_{ij})$ and $\mathcal{X}_i \overset{I[A]}{\Rightarrow}_S \mathcal{X}_j$, we say $\mathcal{X}_i$ indirectly causes $\mathcal{X}_j$ in $S$ and we write $\mathcal{X}_i \overset{I}{\Rightarrow}_S \mathcal{X}_j$; when $A = \text{ind}(\mathcal{I}_{ij})$ and $\mathcal{X}_i \not\overset{I[A]}{\Rightarrow}_S \mathcal{X}_j$, we say that $\mathcal{X}_i$ does not indirectly cause $\mathcal{X}_j$ in $S$ and we write $\mathcal{X}_i \not\overset{I}{\Rightarrow}_S \mathcal{X}_j$.

Again, consider system $S_{18}$ illustrated in $G_{18}$. With $A = \{5, 7\}$, Definition 6.3 states that $\mathcal{X}_2 \overset{I[A]}{\Rightarrow}_S \mathcal{X}_6$ if there exists an admissible intervention $(z_1, z_4, z_6, y_5, r_8(y_7)) \rightarrow (z_2, y_5^*, r_8(y_7^*))$ such that

$$r_9(z_1, z_4, z_6, y_5^*, r_8(y_7^*)) - r_9(z_1, z_4, z_6, y_5, r_8(y_7)) \neq 0.$$

Intuitively, Definition 6.3 concludes that $\mathcal{X}_2 \overset{I[A]}{\Rightarrow}_S \mathcal{X}_6$ if the response of $\mathcal{X}_6$ differs when the effects of setting $\mathcal{X}_2$ to the value $z_2$ as opposed to $z_2^*$ are not transmitted directly, but only through $\mathcal{X}_A$. Thus, setting values for $\mathcal{X}_1$ and $\mathcal{X}_4$ are $z_1$ and $z_4$ in both responses of $\mathcal{X}_6$. On the other hand, setting values for $\mathcal{X}_6$ and $\mathcal{X}_8$ differ across the two responses of $\mathcal{X}_6$ only in response to different settings of $(\mathcal{X}_5, \mathcal{X}_7)$.

When $\mathcal{X}_i \overset{I}{\Rightarrow}_S \mathcal{X}_j$, it follows that for some non-empty $A \subset \mathcal{I}_{ij}$ we have $\mathcal{X}_i \overset{I[A]}{\Rightarrow}_S \mathcal{X}_j$. The converse need not hold, because $\mathcal{X}_i$ can indirectly cause $\mathcal{X}_j$ through each of two distinct intercessors whose associated effects may cancel each other. For example, it may be that $\mathcal{X}_2$ indirectly causes $\mathcal{X}_6$ via $\mathcal{X}_4$ as well as via $\mathcal{X}_6$ in $S_{18}$ but that $\mathcal{X}_2$ does not indirectly cause $\mathcal{X}_6$ via $\{\mathcal{X}_4, \mathcal{X}_6\}$ in $S_{18}$.

6.2.2 Indirect Causality Exclusive of Given Variables

We now introduce an indirect causality concept complementary to that above. For example, in the system illustrated in $G_{16}$, we say that $\mathcal{X}_6$ indirectly causes $\mathcal{X}_3$ exclusive of $\mathcal{X}_i$ if there
exists an admissible intervention \((z_0, z_1, r_2(z_0)) \rightarrow (z_0, z_1, r_2(z_0^*)\)) such that

\[ r_3(z_0, z_1, r_2(z_0^*)) - r_3(z_0, z_1, r_2(z_0)) \neq 0. \]

Similarly, for \(G_{17}\) we say that \(X_0\) indirectly causes \(X_4\) exclusive of \(X_1\) and \(X_3\) if there exists an admissible intervention \((r_2(z_0, z_1)), z_3) \rightarrow (r_2(z_0^*, z_1), z_3)\) such that

\[ r_4(r_2(z_0^*, z_1), z_3) - r_4(r_2(z_0, z_1), z_3) \neq 0. \]

More generally, we say that \(X_i\) indirectly causes \(X_j\) exclusive of \((X_i, X_j)\) —intercessors \(X_A\) if the response of \(X_j\) differs when the effects of setting \(X_i\) to the value \(z_i\) as opposed to \(z_i^*\) are transmitted indirectly through all succeeding variables except \(X_A\).

These examples are instances of the following definition.

**Definition 6.4 Indirect Causality Exclusive of \(X_A\)** Let \(S\) and \(A\) be as Definition 6.3. Then \(X_i\) **indirectly causes \(X_j\) exclusive of \(X_A\) in \(S\)** if there exists an admissible intervention to \((X_{[0:b_1]}(i), X_i, \mathcal{P}_{i,j}, X_A, X_A, \mathcal{X}_X, S^{A}_{i,j})\) with corresponding responses for \(X_j\) such that

\[
\begin{align*}
    r_j(z_{[0:b_1]}(i), z_i, y^*_{i:A}, y^*_{A}, z_A, r_{\mathcal{X}}(z_{[0:b_1]}(i), z_i^*, y^*_{i:A}, z_A), \\
    r_{A:j}^{\mathcal{X}}(z_{[0:b_1]}(i), z_i, y^*_{i:A}, y^*_{A}, z_A, r_{\mathcal{X}}(z_{[0:b_1]}(i), z_i^*, y^*_{i:A}, z_A))) \\
    -r_j(z_{[0:b_1]}(i), z_i, y_{i:A}, y^*_{A}, z_A, r_{\mathcal{X}}(z_{[0:b_1]}(i), z_i, y_{i:A}, z_A), \\
    r_{A:j}^{\mathcal{X}}(z_{[0:b_1]}(i), z_i, y_{i:A}, y^*_{A}, z_A, r_{\mathcal{X}}(z_{[0:b_1]}(i), z_i, y_{i:A}, z_A))) \neq 0;
\end{align*}
\]

and we write \(X_i \overset{I \sim[A]}{\Rightarrow}_S X_j\). Otherwise, we say that \(X_i\) **does not indirectly cause \(X_j\) exclusive of \(X_A\) in \(S\)** and we write \(X_i \not\overset{I \sim[A]}{\Rightarrow}_S X_j\).

In system \(S_{18}\) with \(A = \{5, 7\}\), Definition 6.4 says that \(X_2 \overset{I \sim[A]}{\Rightarrow}_S X_9\) if there exists an admissible intervention \((z_1, y_4, r_6(z_2, z_5), r_8(z_7)) \rightarrow (z_1, y^*_4, r_6(z^*_2, z_5), r_8(z_7))\) such that

\[ r_9(z_1, y^*_4, r_6(z^*_2, z_5), r_8(z_7)) - r_9(z_1, y_4, r_6(z_2, z_5), r_8(z_7)) \neq 0. \]

Intuitively, Definition 6.4 concludes that \(X_2 \overset{I \sim[A]}{\Rightarrow}_S X_9\) if the response of \(X_9\) differs when the effects of setting \(X_2\) to the value \(z_2\) as opposed to \(z^*_2\) are transmitted indirectly through all succeeding variables, except through \(X_A\).
6.3 Total Causality in Recursive Settable Systems

In analyzing relations between causality and conditional independence, it turns out to be important to keep track of channels of both indirect and direct causality. Consider the system illustrated in $G_{15}$ for example. There, we say that $X_0$ (totally) causes $X_1$ if there exists an admissible intervention $(z_0, r_1(z_0)) \rightarrow (z_0^*, r_1(z_0^*))$ such that

$$r_2(z_0^*, r_1(z_0^*)) - r_2(z_0, r_1(z_0)) \neq 0.$$ 

Intuitively, the response of $X_2$ differs when the effect of setting $Z_0$ to the value $z_0$ as opposed to $z_0^*$ is transmitted fully, taking into account both direct and indirect effects. Similarly, in the system illustrated in $G_{16}$ we say that $X_0$ (totally) causes $X_1$ if there exists an admissible intervention $(z_0, r_1(z_0), z_2) \rightarrow (z_0^*, r_1(z_0^*), z_2)$ such that

$$r_2(z_0^*, r_1(z_0^*), z_2) - r_2(z_0, r_1(z_0), z_2) \neq 0.$$ 

We now provide formal definitions of (total) causality via and exclusive of a set of variables.

**Definition 6.5 A–Causality** Let $S$ and $A$ be as Definition 6.3. Then $X_i$ **causes** $X_j$ via $X_A$ (or $X_i$ **A–causes** $X_j$) in $S$ if there exists an admissible intervention to $(X_{[0:b_i]}(i), X_i, P^A_{i:j}, X_A, X_{\bar{A}}, S^{i:j})$ with corresponding responses for $X_j$ such that

$$r_j(z_{[0:b_i]}(i), z^*_i, z_{i:A}, z_{\bar{A}}, y^*_A, r_{\bar{A}}(z_0, z_{[1:b_i]}(i), z_i), z_{i:A}, y^*_A),$$

$$r_{\bar{A}:j}[z_{[0:b_i]}(i), z_i, z_{i:A}, z_{\bar{A}}, y^*_A, r_{\bar{A}}(z_0, z_{[1:b_i]}(i), z_i), z_{i:A}, y^*_A]) \neq 0;$$

and we write $X_i \Rightarrow^A_S X_j$. Otherwise, we say that $X_i$ **does not A–cause** $X_j$ in $S$ and we write $X_i \not\Rightarrow^A_S X_j$. When $A = \text{ind}(I_{i:j})$ and $X_i \Rightarrow^A_S X_j$, we say $X_i$ **causes** $X_j$ in $S$ and we write $X_i \Rightarrow_S X_j$; when $A = \text{ind}(I_{i:j})$ and $X_i \not\Rightarrow^A_S X_j$, we say that $X_i$ **does not cause** $X_j$ in $S$ and we write $X_i \not\Rightarrow_S X_j$.

**Definition 6.6 ~ A–Causality** Let $S$ and $A$ be as Definition 6.3. Then $X_i$ **causes** $X_j$ **exclusive of** $X_A$ (or $X_i \sim A–causes$ $X_j$) in $S$ if there exists an admissible intervention to $(X_{[0:b_i]}(i), X_i, P^A_{i:j}, X_A, X_{\bar{A}}, S^{i:j})$ with corresponding responses for $X_j$ such that

$$r_j(z_{[0:b_i]}(i), z^*_i, y^*_A, z_A, r_{\bar{A}}(z_0, z_{[1:b_i]}(i), z_i), y_{i:A}, z_A),$$

$$r_{\bar{A}:j}[z_{[0:b_i]}(i), z_i, y^*_A, z_A, r_{\bar{A}}(z_0, z_{[1:b_i]}(i), z_i), y_{i:A}, z_A]) \neq 0;$$

$$r_{\bar{A}:j}[z_{[0:b_i]}(i), z_i, y^*_A, z_A, r_{\bar{A}}(z_0, z_{[1:b_i]}(i), z_i), y_{i:A}, z_A]) \neq 0.$$
and we write $\mathcal{X}_i \Rightarrow^A_S \mathcal{X}_j$. Otherwise, we say that $\mathcal{X}_i$ \textit{does not cause} $\mathcal{X}_j$ \textit{exclusive of} $\mathcal{X}_A$ (or $\mathcal{X}_i \not\sim A$–\textit{cause} $\mathcal{X}_j$) in $S$, and we write $\mathcal{X}_i \not\Rightarrow^A_S \mathcal{X}_j$.

Thus, Definitions 6.5 and 6.6 are analogous to Definitions 6.3 and 6.4 with the difference that the direct effect of $\mathcal{X}_i$ on $\mathcal{X}_j$ is now further taken into account.

6.4 Relations among Total, Direct, and Indirect Causality

We now relate the various causality notions defined above. These relations are completely intuitive, but it is important that they be made rigorous. Moreover, their plausibility suggests that the foregoing definitions are natural in an important sense.

First, we link $A$–causality, direct causality, and indirect causality via $\mathcal{X}_A$. This relates to some useful basic results on (indirect) causality via or exclusive of $\mathcal{X}_A$ for the special cases $A = \emptyset$ or $A = \text{ind}(\mathcal{I}_{i,j})$, given in Proposition 10.1 of the appendix.

**Proposition 6.1** Let $S$ and $A$ be as Definition 6.3 and suppose that $\mathcal{X}_i \Rightarrow^A_S \mathcal{X}_j$. Then $\mathcal{X}_i \Rightarrow^D_S \mathcal{X}_j$ or $\mathcal{X}_i \Rightarrow^I_S \mathcal{X}_j$ or both.

An important special case of Proposition 6.1 occurs when $A = \text{ind}(\mathcal{I}_{i,j})$.

**Corollary 6.2** Let $S$ and $A$ be as Definition 6.3 and suppose that $\mathcal{X}_i \Rightarrow_S \mathcal{X}_j$. Then $\mathcal{X}_i \Rightarrow^D_S \mathcal{X}_j$ or $\mathcal{X}_i \Rightarrow^I_S \mathcal{X}_j$ or both.

Corollary 6.2 verifies the plausible claim that if $\mathcal{X}_i$ causes $\mathcal{X}_j$, it does so directly, indirectly, or both. Significantly, the converse need not hold, as direct and indirect causal channels can cancel one another. Proposition 6.1 extends this proposition to $A$–causality. A similar result holds for $\not\sim A$–causality:

**Proposition 6.3** Let $S$ and $A$ be as Definition 6.3, and suppose that $\mathcal{X}_i \Rightarrow^A_S \mathcal{X}_j$. Then $\mathcal{X}_i \Rightarrow^D_S \mathcal{X}_j$ or $\mathcal{X}_i \Rightarrow^I_S \mathcal{X}_j$ or both.

It is possible and of interest to study even more refined notions of (indirect) causality using this framework. For example, for disjoint subsets $A$ and $B$ of $\text{ind}(\mathcal{I}_{i,j})$, we can study the notions of $\mathcal{X}_i$ (indirectly) causing $\mathcal{X}_j$ (a) via $A$ and via $B$; (b) via $A$ and exclusive of $B$; (c) via $A$ or exclusive of $B$; and (d) exclusive of $A$ or exclusive of $B$. For brevity, we leave a formal treatment of these causal notions to other work.
7 Conditional Independence in Recursive Systems

In this section, we begin formal study of the connections between conditional independence and the settable system notions of causality introduced in Section 6. Our first result relates causality and conditional dependence by establishing the conditional Reichenbach principle. This implies the traditional unconditional Reichenbach principle. We then extend conditional Reichenbach to give necessary and sufficient conditions relating conditional independence and causality in recursive settable systems. Among other things, this enables us to relate graphical separation and causality, taken up in the next section.

We focus on canonical systems. Recall that the canonical setting for $\mathcal{X}_i, i \in \Pi_b,$ is

$$Z_i^c = Y_i := r_i(Z_{[0:b-1]}).$$

Letting this expression recursively define the canonical settings $Z_{[0:b-1]}^c, b = 1, ..., B,$ with $Z_0^c := Z_0 = Y_0 =: Y_0^c,$ we also define canonical responses

$$Y_i^c := r_i(Z_{[0:b-1]}^c), \quad i \in \Pi_b, \quad b = 1, ..., B.$$  

In what follows, whenever we reference canonical responses, we implicitly assume their existence.

7.1 The Conditional Reichenbach Principle of Common Cause

So far, none of our definitions or results have required any probabilistic elements. To relate causality and probabilistic dependence, we now explicitly introduce probability measures $P$ on $(\Omega, \mathcal{F})$. Our next result formalizes a conditional version of Reichenbach’s principle.

Proposition 7.1 The Conditional Reichenbach Principle of Common Cause (I)

Let $S$ be a recursive partitioned settable system, and for $a, b \geq 0,$ let $i \in \Pi_a$ and $j \in \Pi_b,$ $i \neq j$. Let $\mathcal{X}_i$ and $\mathcal{X}_j$ be settable variables with canonical responses $Y_i^c$ and $Y_j^c$. Let $A \subset \Pi \setminus \{i, j\},$ and let $\mathcal{X}_A$ be the corresponding vector of settable variables with canonical responses $Y_A^c$. For every probability measure $P$ on $(\Omega, \mathcal{F}),$ if $Y_i^c \not\perp Y_j^c \mid Y_A^c$ then either:

(i) $i = 0$ and $\mathcal{X}_0$ causes $\mathcal{X}_j$ exclusive of $A_j := A \cap \text{ind}(\mathcal{I}_{0:j}),$ i.e., $\mathcal{X}_0 \Rightarrow_{A_j} \mathcal{X}_j$; or

(ii) $j = 0$ and $\mathcal{X}_0$ causes $\mathcal{X}_i$ exclusive of $A_i := A \cap \text{ind}(\mathcal{I}_{0:i}),$ i.e., $\mathcal{X}_0 \Rightarrow_{A_i} \mathcal{X}_i$; or

(iii) $i, j \neq 0$ and $\mathcal{X}_0 \Rightarrow_{A_j} \mathcal{X}_j$ and $\mathcal{X}_0 \Rightarrow_{A_i} \mathcal{X}_i$.

The traditional Reichenbach principle of common cause follows by putting $A = \emptyset$. 

36
Corollary 7.2 The Reichenbach Principle of Common Cause

Let $S$, $X_i$, and $X_j$ be as in Proposition 7.1. For every probability measure $P$ on $(\Omega, \mathcal{F})$, if $Y_i^c \perp Y_j^c$, then either:

(i) $i = 0$ and $X_0$ causes $X_j$, i.e., $X_0 \Rightarrow_S X_j$; or

(ii) $j = 0$ and $X_0$ causes $X_i$, i.e., $X_0 \Rightarrow_S X_i$; or

(iii) $i, j \neq 0$ and $X_0 \Rightarrow_S X_i$ and $X_0 \Rightarrow_S X_j$.

This provides fully explicit conditions, both causal and probabilistic, under which the Reichenbach principle of common cause holds – that is, under which it is true that when canonical responses for two settable variables are probabilistically dependent, either one causes the other or there exists an underlying common cause. Note that while the possibility that one variable causes the other is not explicit in (iii), it is nevertheless implicit, as one way in which we may have $X_0 \Rightarrow_S X_j$ is via the indirect channel $X_0 \Rightarrow_S X_i \Rightarrow_S X_j$. If this fails in (iii), then there nevertheless must be a common cause, $X_0$.

This analysis reveals that the traditional unconditional Reichenbach principle is not a deep fact. The reason is that the principal settable variable $X_0$ can always serve as a universal common cause. Moreover, because the principal setting values $z_0$ are identified with the underlying elements $\omega_0$ of the principal universe $\Omega_0$, one cannot dispense with this universal common cause without dispensing with the underlying structure supporting probability statements. This demonstrates the indispensable and dramatically simplifying role played by the principal variable $X_0$ as a universal common cause. Once this role is understood, the content of the unconditional Reichenbach principle is no longer mysterious. Its previously perplexing status can be understood as a consequence of the lack of a proper context for its formulation. The settable system framework supplies this context.

The conditional Reichenbach principle is substantive, however, as it implies that in recursive causal systems, knowledge of conditional dependence relations such as $Y_i^c \perp Y_j^c \mid Y_A^c$ is informative about the possible causal relations holding between settable variables $X_i$ and $X_j$. Proposition 7.1 implies that in recursive systems, in order for two canonical responses $Y_i^c$ and $Y_j^c$ to be conditionally dependent given a vector of canonical responses $Y_A^c$, it must be that the principal variable $X_0$ causes at least $X_i$ or $X_j$ exclusive of the relevant subsets of $X_A$. Otherwise, we can express $Y_i^c$ or $Y_j^c$ (or both) as a function of the relevant sub-vector of $Y_A^c$. As Proposition 7.1 has $A \subset \Pi \setminus \{i, j\}$, it is necessary that $0 \notin A$; $Y_i^c \perp Y_j^c \mid Y_A^c$ cannot hold otherwise.

Further, the possibility that one variable causes the other is again implicit in (iii): one way in which we may have $X_0 \Rightarrow_{A_i} X_j$ and $X_0 \Rightarrow_{A_j} X_i$ is via the indirect channel $X_0 \Rightarrow_{A_i} X_i \Rightarrow_{A_{ij}} X_j$ with $A_{ij} := A \cap \text{ind}(I_{ij})$. But even if this fails in (iii), then there nevertheless must be a common cause, $X_0$. 

37
It is a useful fact that if the conclusion of Proposition 7.1 holds (regardless of whether the stated conditions hold) then the direct causality graph \( G \) associated with a system \( S \) has the following simple property.

**Proposition 7.3** Let \( S, \mathcal{X}_i, \mathcal{X}_j, \) and \( \mathcal{X}_A \) be as in Proposition 7.1, and let \( G \) be the associated direct causality graph. Suppose that conclusions (i) – (iii) of Proposition 7.1 hold. Then there exist an \((\mathcal{X}_0, \mathcal{X}_i)\) path (if \( i \neq 0 \)) and an \((\mathcal{X}_0, \mathcal{X}_j)\) path (if \( j \neq 0 \)) that does not contain elements of \( \mathcal{X}_A \).

Thus, it suffices for \( Y_i^c \perp Y_j^c \mid Y_A^c \) that \( \mathcal{X}_0 \not\sim_A \mathcal{X}_i \) or \( \mathcal{X}_0 \not\sim_A \mathcal{X}_j \), which in turn is implied by the absence of an \((\mathcal{X}_0, \mathcal{X}_i)\) path or an \((\mathcal{X}_0, \mathcal{X}_j)\) path that does not contain elements of \( \mathcal{X}_A \).

To illustrate, we apply Proposition 7.3 to system \( S_{18} \). We have \( Y_0^c \perp Y_i^c \mid Y_2^c \) for \( i = 3, ..., 8 \), as \( \mathcal{X}_0 \not\sim^{(2)} \mathcal{X}_i \) for \( i = 3, ..., 8 \). Similarly, \( Y_0^c \perp Y_9^c \mid (Y_1^c, Y_2^c) \), as \( \mathcal{X}_0 \not\sim^{(1,2)} \mathcal{X}_9 \).

Also, we have that \( Y_2^c \perp Y_5^c \mid Y_3^c \), since \( \mathcal{X}_0 \not\sim \mathcal{X}_5 \). In addition, we have that \( Y_3^c \perp Y_5^c \mid Y_2^c \), as \( \mathcal{X}_0 \not\sim \mathcal{X}_3 \) and \( \mathcal{X}_0 \not\sim \mathcal{X}_5 \). These facts hold for every \( P \).

These examples only require knowledge of direct causality relations. The direct causality graph suffices for this. But specific properties of the response functions, not indicated by the graph, may also be important. To illustrate, consider determining whether \( Y_2^c \perp Y_3^c \) in \( S_{18} \). Corollary 7.2 gives that either \( \mathcal{X}_0 \not\sim \mathcal{X}_2 \) or \( \mathcal{X}_0 \not\sim \mathcal{X}_3 \) (or both) is sufficient for this to hold. We know from \( G_{18} \) that \( \mathcal{X}_0 \rightarrow \mathcal{X}_2 \), but determining whether \( \mathcal{X}_0 \not\sim \mathcal{X}_3 \) requires additional information about the functional form of response functions \( r_2 \) and \( r_3 \).

Similarly, consider whether \( Y_2^c \perp Y_7^c \mid Y_3^c \) holds in \( S_{18} \). From the contrapositive of Proposition 7.1 we know that this will hold if \( \mathcal{X}_0 \not\sim \mathcal{X}_7 \). Nevertheless, determining whether \( \mathcal{X}_0 \not\sim \mathcal{X}_7 \) holds requires additional information about the functional forms of the response functions not contained in \( G_{18} \).

Thus, similar to the situation for PCM DAGs, settable system direct causality graphs do not provide complete information about the causal relations needed for resolving questions of conditional independence. Indeed, as argued in Dawid (2010a), nongraphical representations of causality are indispensable here. Our function-based definitions of causality supply just the information needed to relate causality to conditional independence.

Because a direct causality graph is not a probabilistic DAG, there is no reason to expect \( d \)-separation to be informative about conditional independence in direct causality DAGs, such as \( G_{18} \). For example, although \( Y_3^c \perp Y_5^c \mid Y_2^c \), \( \mathcal{X}_3 \) and \( \mathcal{X}_5 \) are not \( d \)-separated by \( \mathcal{X}_2 \) in
Similarly, we have $Y^c_2 \perp Y^c_5 \mid (Y^c_3, Y^c_6)$, since $X_0 \not\perp^S X_3$, whereas $X_2$ and $X_5$ are not $d$-separated by $(X_3, X_6)$ in $G_{18}$, due to the “collider” $X_2 \rightarrow X_6 \leftarrow X_3$. There is no paradox here: $d$-separation may imply conditional independence in a certain class of probabilistic DAGs, but it does not generally apply to direct causality graphs.

Note also that because the path \{\x_2, \x_6, \x_7\} does not contain \x_3, we have that \x_2 and \x_7 are not $d$–separated by \x_3 in $G_{18}$. To conclude that $Y^c_2 \not\perp Y^c_3$ and $Y^c_2 \not\perp Y^c_6 \mid Y^c_3$ in such situations in PCM DAGs, Pearl (2000, p. 48-49) and SGS (pp. 35, 56) introduce the assumptions of “stability” or “faithfulness” of the probability measure $P$. In sharp contrast, Proposition 7.1 imposes no restrictions on $P$; instead the properties of the response functions play the key role in determining the presence or absence of causal relations.

### 7.2 Characterizing the Conditional Reichenbach Principle

The conditional Reichenbach principle gives necessary but not sufficient causal conditions for conditional dependence. Thus, its contrapositive gives sufficient but not necessary conditions for conditional independence. Specifically, $Y_i^c \perp Y_j^c \mid Y_k^c$ can hold even when the conclusion of Proposition 7.1 holds. Examples of this are easy to construct.

**Example 7.4** Consider system $S_{16}$ in $G_{16}$ and suppose that $Y_i^c$ and $Y_2^c$ are jointly normal with mean zero, variance one, and correlation $\rho$. Then $Y_i^c$ and $Y_2^c$ are independent if and only if $\rho = 0$. When $\rho = 0$, $Y_i^c \perp Y_2^c$ even though $\x_1$ and $\x_2$ share the common cause $\x_0$.

It is also easy to construct examples in which independence holds between directly causally related variables.

**Example 7.5** Consider system $S_{16}$ in $G_{16}$ and suppose that $Y_i^c$ and $Y_2^c$ are jointly normal with mean zero, variance one, and correlation $\rho$. Suppose also that $\x_0 \not\perp^S \x_3$, with

$$Y_3^c = Y_1^c + aY_2^c.$$  

Then $Y_2^c$ and $Y_3^c$ are also jointly normal, with mean zero. Let $a = -\rho$. Then $Y_3^c$ and $Y_2^c$ have zero correlation, so they are independent, even though $\x_2 \not\perp^D \x_3$. (Note that $Y_3^c$ has non-zero variance as long as $|a| < 1$.)

It is thus useful to refine the possibilities for conditional independence to distinguish (a) situations in which causal restrictions among settable variables ensure that their canonical responses are conditionally independent for any probability measure and (b) those where
conditional independence holds only for *some* choice of $P$. Direct causality restrictions may be sufficient but are not necessary for (a) to obtain, as seen above. Also, (b) can hold due to: (i) a particular choice of $P$ only; or (ii) both a particular configuration of response functions and a particular choice of $P$. The following definitions are useful for this.

**Definition 7.1** Conditional Causal Isolation and Conditional $P$-Stochastic Isolation

Let $S$, $Y^c_i$, $Y^c_j$, and $Y^c_A$ be as in Proposition 7.1. Suppose that the conclusion of Proposition 7.1 fails; then $\mathcal{X}_i$ and $\mathcal{X}_j$ are causally isolated given $\mathcal{X}_A$. Let $P$ be a probability measure on $(\Omega, \mathcal{F})$ and suppose that $Y^c_i \perp Y^c_j | Y^c_A$ when $\mathcal{X}_i$ and $\mathcal{X}_j$ are not causally isolated given $\mathcal{X}_A$; then we say that $\mathcal{X}_i$ and $\mathcal{X}_j$ are $P$-stochastically isolated given $\mathcal{X}_A$.

From Definition 7.1, we have that $\mathcal{X}_i$ and $\mathcal{X}_j$ are causally isolated given $\mathcal{X}_A$ when $\mathcal{X}_0 \not\equiv_S \mathcal{X}_i$ or $\mathcal{X}_0 \not\equiv_S \mathcal{X}_j$. The “isolation” is from the potential cause $\mathcal{X}_0$. When $A = \emptyset$, we say that $\mathcal{X}_i$ and $\mathcal{X}_j$ are causally isolated when the conclusion of Corollary 7.2 does not hold, that is, when $\mathcal{X}_0 \not\equiv_S \mathcal{X}_i$ or $\mathcal{X}_0 \not\equiv_S \mathcal{X}_j$. Conditional causal isolation arises when, for one or the other of $\mathcal{X}_i$ and $\mathcal{X}_j$, the response functions channel the effects of the principal cause $\mathcal{X}_0$ in just the right way so as to yield canonical responses $Y^c_i$ or $Y^c_j$ (or both) expressible just as a function of the relevant subsets of $Y^c_A$ (i.e., $Y^c_{A_i}$ or $Y^c_{A_j}$).

Conditional $P$-stochastic isolation is just conditional independence without conditional causal isolation. It can arise either from $P$ alone, as in Example 7.4, or from just the right combination of $P$ and functional relations between multiple causes (common or direct), as in Example 7.5. In fact, if the conditional distributions of $Y^c_i$ and $Y^c_j$ given $Y^c_A$ are each regular (see e.g. Dudley, 2002, p. 341–344), then there is always a joint probability measure $P^*$ ensuring that $Y^c_i$ and $Y^c_j$ are conditionally independent given $Y^c_A$, regardless of the causal relations involving $\mathcal{X}_i$, $\mathcal{X}_j$, and $\mathcal{X}_A$ (see proposition III.2.1 of Neveu, 1965, p. 74-75). $P$-stochastic isolation is, however, a nontrivial restriction. Thus, the utility of this concept is that it permits us to distinguish between guaranteed sources of conditional independence (conditional causal isolation) and more special or exceptional cases.

For the same reason, conditional $P$-stochastic isolation should not be casually assumed. Instead, it should be empirically subjected to falsification whenever feasible, by testing the conditional independence(s) it may be thought to justify. See White and Chalak (2010) and the references given there for results delivering empirical tests of conditional independence.

We can now characterize the relation between causality and conditional independence for canonical responses $Y^c_i$ and $Y^c_j$ in recursive settable systems given any vector of canonical responses $Y^c_A$. For clarity, we state this in the contrapositive.
Corollary 7.6 Conditional Reichenbach Principle of Common Cause (II) Suppose the conditions of Proposition 7.1 hold. For given probability measure $P$ on $(\Omega, \mathcal{F})$, $Y_i^c \perp Y_j^c \mid Y_k^c$ if and only if either (a) $X_i$ and $X_j$ are causally isolated given $X_A$; or (b) $X_i$ and $X_j$ are $P$-stochastically isolated given $X_A$.

When $A = \emptyset$, Corollary 7.6 strengthens Reichenbach’s principle of common cause to give necessary and sufficient causal conditions for probabilistic dependence. For the empirically relevant case where the canonical responses may all be vectors, Theorem 10.4 of the appendix formally characterizes the relations between conditional independence and causality.

8 Settable Systems and Graphical Separation

As we discuss in Section 2, implications of $d$-separation in probabilistic DAGs have sometimes been ascribed causal intuition (e.g., Pearl, 2000, p. 16-17). Absent other causal relations and expressed in the present notation and nomenclature, these can be stated for canonical responses $Y_i^c$ and $Y_j^c$ as:

- $d.1 \ Y_i^c \perp Y_j^c \mid Y_k^c$, provided $X_A$ fully mediates the effect of $X_i$ on $X_j$;
- $d.2 \ Y_i^c \perp Y_j^c \mid Y_k^c$, provided $X_A$ denotes the common causes for $X_i$ and $X_j$, or fully mediates the effects of these common causes on either (or both) $X_i$ or $X_j$;
- $d.3 \ Y_i^c \nmid Y_j^c \mid Y_k^c$, provided $X_A$ is caused by both $X_i$ and $X_j$.

We reiterate that in the context of probabilistic DAGs, such causal interpretations are problematic. In contrast, causal semantics are well defined in settable systems, allowing the truth values of $d.1 \text{--} d.3$ to be assessed. As the examples of Sections 3 and 7 show, $d.1$ and $d.2$ may or may not hold. Similarly, $d.3$ may or may not hold. In this section, we provide conditions ensuring that these statements do indeed hold, paying particular attention to the strength of the conditions required to ensure each property.

First, we consider a quite special subclass of settable systems, directly analogous to Markovian and semi-Markovian PCMs, in which the directed local Markov property and hence $d$-separation holds, so that $d.1$ and $d.2$ hold. Next, we discuss other quite special settable systems where conditional independence relations additional to the local Markov property may hold, as encoded by the $D$-separation criteria discussed in Geiger et. al. (1990). This ensures $d.1$ and $d.2$, as well as other statements that generally fail in Markovian systems. Finally, we provide an extended version of $d.3$ for settable systems. The conditions for this are fairly general. In particular, neither the local Markov property nor notions of stability or faithfulness are required.
Notions of $d$-separation and $D$-separation, as well as their underlying assumptions, are therefore not fundamental to establishing the connections between functionally defined causal relations and conditional independence, nor are they a natural starting point or context for this study. Nevertheless, as we show, they can be helpful for verifying or falsifying conditional independence relations in suitably restricted settable systems.

In this section, we will always reference canonical responses. Accordingly, we drop the explicit superscript “$\circ$” and write $Y_j$ in place of $Y_j^\circ$, etc., for notational convenience.

8.1 Conditioning on Predecessors

8.1.1 The Markovian PCM and $d$–Separation in Settable Systems

Although $d.1$ and $d.2$ do not hold in settable systems generally, we now provide conditions under which they are true. Our next result describes a settable system analogous to the Markovian PCM that generalizes the examples illustrated in $G_0$ through $G_{13}$. In particular, we show that here the local Markov property holds for certain random variables analogous to endogenous variables in the PCM.

**Proposition 8.1** Let $S$ be recursive. Suppose that $\mathcal{X}_0 \xrightarrow{D} \mathcal{X}_k$ for all $k \in \Pi_1$ and that the elements of $\Pi_1$ are in one-to-one correspondence with those of $\Pi_{[2:B]}$, such that for each $i \in \Pi_{[2:B]}$, there is a unique $k \in \Pi_1$ such that $\mathcal{X}_k \xrightarrow{D} \mathcal{X}_i$. Suppose further that $\mathcal{X}_0 \not\xrightarrow{D} \mathcal{X}_i$ for all $i \in \Pi_{[2:B]}$. For given $i \in \Pi_b$, $b \geq 2$, let $C := \{l \in \Pi_{[2:b-1]} : \mathcal{X}_l \xrightarrow{D} \mathcal{X}_i\}$ and let $A := \{j \in \Pi_{[2:B]} \setminus C : \mathcal{X}_j$ does not succeed $\mathcal{X}_i\}$. Let $Y_i, Y_C$, and $Y_A$ be canonical responses of $\mathcal{X}_i$, $\mathcal{X}_C$, and $\mathcal{X}_A$. If $P$ is a probability measure on $(\Omega, \mathcal{F})$ such that $\{Y_k : k \in \Pi_1\}$ are jointly independent, then $Y_i^\circ \perp Y_A^\circ \mid Y_C^\circ$.

A special case of Proposition 8.1 obtains for $C = \emptyset$, in which case $Y_i \perp Y_A$ follows.

Here, conditional independence holds but not conditional causal isolation, an example of $P$–stochastic causal isolation. Such systems are very special indeed, in that probability measures $P$ ensuring the joint independence of $\{Y_k : k \in \Pi_1\}$ are shy in the set of all joint probability distributions.\(^6\) Shyness is the function space analog of being a subset of a set of Lebesgue measure zero.

It is easy to construct a probabilistic DAG compatible with the distribution of canonical responses $\{Y_i : i \in \Pi_{[2:B]}\}$. This DAG is isomorphic to the subgraph of the settable system direct causality graph corresponding to elements of $\Pi_{[2:B]}$, substituting canonical responses for settable variables at the nodes. Further, Lauritzen et. al. (1990, proposition 3) ensures

\(^6\)For a discussion of shyness, see Corbe, et. al. (2009).
that for such systems, \(d\)-separation or equivalent graphical criteria can identify exactly the conditional independence relations implied by the directed local Markov property.

For example, consider the agent/expert example illustrated in direct causality graph \(G_8\), where the expert’s advice does not directly affect the outcome. With no restrictions on \(P\), the expert’s advice, \(Y_4\), and the outcome, \(Y_6\), need not be conditionally independent given the canonical agent action, \(Y_5\), since \(X_4\) and \(X_6\) need be not causally isolated given \(X_5\). This is despite the fact that the agent’s action \(X_5\) fully mediates the effect of the advice \(X_4\) on the outcome \(X_6\). Nevertheless, if we impose the strong assumption that the causal structure is as depicted in \(G_8\) with \((Y_1, Y_2, Y_3)\) jointly independent, then Proposition 8.1 ensures \(Y_4 \perp Y_6 \mid Y_5\), so that \(X_4\) and \(X_6\) are \(P\)-stochastically isolated given \(X_5\). This is illustrated in the probabilistic DAG \(G_9\) associated with this “Markovian” structure. There, \(Y_5\) \(d\)-separates \(Y_4\) and \(Y_6\).

### 8.1.2 Deterministic and Chance Nodes and \(D\)-separation

Geiger et. al. (1990) study DAGs that distinguish between “deterministic” and “chance” nodes. A deterministic node corresponds to a random variable that is conditionally independent of all other random variables given its DAG parents, whereas a chance node corresponds to a random variable that is conditionally independent of its “non-descendants” (non-successors) given its parents.

Geiger et. al. (1990) call the conditional independence statements corresponding to deterministic and chance nodes an “enhanced basis” and provide an analogue to \(d\)-separation for these DAGs called “\(D\)-separation” that ensures conditional independence under the graphoid axioms. Similar to probabilistic DAGs, these DAGs do not contain any necessary causal content. In traditional Markovian PCM graphs (such as \(G_9\)), none of the nodes are fully determined by their parents, so it follows that \(d\)-separation and \(D\)-separation coincide in such DAGs.

Suitably restricted settable systems can embody \(D\)-separation. For example, consider the probabilistic DAG \(G^*\) corresponding to the direct causality graph \(G\) for a recursive system \(S\) that substitutes canonical responses for settable variables. Theorem 10.4 says that if \(C \subset \text{ind}(I_{0:i})\) is such that \(X_0 \not\sim^C_S X_i\) and \(A = (\Pi \cup \Pi_0) \setminus \{\{i\} \cup C\}\), then \(Y_i \perp Y_A \mid Y_C\).

In particular, \(X_0 \not\sim^C_S X_i\) holds when the set \(C\) corresponds to all direct causes (i.e., the “parents”) of \(X_i\). In this sense, \(Y_0\) is a chance node in \(G^*\), whereas the nodes \(Y_i, i \neq 0\), in \(G^*\) are deterministic.

It can be verified that for disjoint sets \(D, E, F\) in \(\Pi \cup \Pi_0\), \(Y_D\) and \(Y_E\) are not \(D\)-separated given \(Y_F\) in \(G^*\) if and only if: (a) \((i) 0 \in D\) and (ii) for some \(j \in E\), there
exists an \((\mathcal{X}_0, \mathcal{X}_j)\) path in \(G\) that does not contain elements of \(\mathcal{X}_F\); or (b)(i) \(0 \in E\) and (ii) for some \(i \in D\), there exists an \((\mathcal{X}_0, \mathcal{X}_i)\) path in \(G\) that does not contain elements of \(\mathcal{X}_F\); or (c)(i) \(0 \notin D \cup E\) and (ii) (a.ii) and (b.ii) hold.

Although the graphical \(D\)-separation criteria are sufficient for \(Y_D \perp Y_E \mid Y_F\), they are not necessary. A more general sufficient condition for \(Y_D \perp Y_E \mid Y_F\) is the failure of Theorem 10.4’s condition (a), as this is implied by, but does not imply, \(D\)-separation in \(G^*\). This is a further example of the limitations of graphical criteria.

Our next result describes a restricted settable system similar to that in Proposition 8.1 that generates random variables forming an enhanced basis.

**Proposition 8.2** Let \(S\) be recursive. Suppose that \(\mathcal{X}_0 \overset{D}{\rightarrow} \mathcal{X}_k\) for all \(k \in \Pi_1\) and that for each \(k \in \Pi_1\), there is a unique \(i \in \Pi_{[2:B]}\) such that \(\mathcal{X}_k \overset{D}{\rightarrow} \mathcal{X}_i\). Suppose further that \(\mathcal{X}_0 \not\overset{D}{\rightarrow} \mathcal{X}_i\) for all \(i \in \Pi_{[2:B]}\). For given \(i \in \Pi_b, b \geq 2\), let \(C := \{l \in \Pi_{[2:b-1]} : \mathcal{X}_l \overset{D}{\rightarrow} \mathcal{X}_i\}\). Let \(A_1 := \{l \in \Pi_{[2:B]\setminus(C \cup \{i\})} : \mathcal{X}_j \text{ does not succeed } \mathcal{X}_i\}\). Let \(P\) be a probability measure on \((\Omega, \mathcal{F})\).

(i) Suppose that \(\mathcal{X}_k \not\overset{D}{\rightarrow} \mathcal{X}_i\) for all \(k \in \Pi_1\). Then \(Y_i \perp Y_{A_1} \mid Y_C\).

(ii) Suppose that (a) \(\mathcal{X}_k \overset{D}{\rightarrow} \mathcal{X}_i\) for some \(k \in \Pi_1\) and (b) \(P\) is such that \(\{Y_k : k \in \Pi_1\}\) are jointly independent. Then \(Y_i \perp Y_{A_2} \mid Y_C\).

Part (i) ensures conditional causal isolation. Part (ii) gives \(P\)-stochastic isolation when elements of \(\{Y_k : k \in \Pi_1\}\) are jointly independent, generating an enhanced basis involving \(\{Y_i, i \in \Pi_{[2:B]}\}\). This is represented in the probabilistic \(\text{DAG } G^\dagger\) isomorphic to the subgraph for \(i \in \Pi_{[2:B]}\) of \(G\), substituting canonical responses for settable variables at the nodes. If \(\mathcal{X}_k \not\overset{D}{\rightarrow} \mathcal{X}_i\) for all \(k \in \Pi_1\), then \(Y_i\) is represented by a (dashed) deterministic node in \(G^\dagger\). Otherwise, \(Y_i\) is represented by a (solid) chance node. Applying the \(D\)-separation criteria to \(G^\dagger\) identifies exactly the conditional independence relations implied by this enhanced basis under the graphoid axioms.

To illustrate, consider the canonical responses of the expert/agent example illustrated in direct causality graph \(G_{19}\) where the expert’s advice, \(\mathcal{X}_3\), has no effect on the outcome, \(\mathcal{X}_5\), and the agent fully complies with the expert’s advice.
Since $\mathcal{X}_0 \not \sim^3 \mathcal{X}_4$ in $\mathcal{S}_{19}$, Lemma 10.2 ensures that the agent’s action canonical response $Y_4$ is determined as a function of the expert’s advice canonical response $Y_3$. Thus, $Y_4$ is represented by a deterministic (dashed) node in probabilistic DAG $G_{20}$. On the other hand, $Y_3$ and $Y_5$ are represented by chance (solid) nodes in $G_{20}$. Proposition 8.2 gives that $Y_4 \perp Y_5 \mid Y_3$ for any $P$, a consequence of conditional causal isolation. If $Y_1$ and $Y_2$ are independent, then $\mathcal{X}_3$ and $\mathcal{X}_5$ are $P$-stochastically isolated given $\mathcal{X}_4$ and $Y_3 \perp Y_5 \mid Y_4$.

Such structures impose very strong restrictions on both the causal relationships of the system and on the distribution of the responses in the first block, $\{Y_k : k \in \Pi_1\}$.

### 8.2 Conditioning on Successors

Unlike properties $d.1$ and $d.2$, which concern the independence of successors conditioning on predecessors, $d.3$ is a statement about the (lack of) conditional independence of predecessors, conditioning on successors. Without further conditions, there is no guarantee that $d.3$ holds in recursive settable systems. For example, for canonical responses in system $\mathcal{S}_{18}$, we have that $Y_2 \perp Y_5 \mid (Y_3, Y_6)$ since $\mathcal{X}_0 \not \sim^3 \mathcal{X}_5$ even though $\mathcal{X}_2 \not \perp \mathcal{X}_6$ and $\mathcal{X}_5 \perp \mathcal{X}_6$.

Nevertheless, $d.3$ may follow from the conditional Reichenbach principle for special cases under further assumptions. For instance, the local Markov property holds in Markovian systems as described in Section 8.2; thus, $d.3$ holds in such systems provided faithfulness (SGS, pp. 35, 56) or stability (Pearl 2000, p. 48-49) holds. But these notions do not provide satisfying insight, as they amount to assuming whatever might be needed to ensure that $d.3$ holds. Moreover, the local Markov property is quite a strong restriction.

Our next result gives general conditions ensuring that an extension of $d.3$ holds for systems that need not be Markovian (see also Wermuth and Cox, 2004, section 7). We extend $d.3$ by allowing conditioning on both successors and non-successors.
Theorem 8.3 Let $S$ be recursive. Let $A, B, C,$ and $D$ be disjoint subsets of $\Pi$ such that for all $i \in D$ and $j \in A \cup B \cup C$, $X_i$ does not precede $X_j$. Suppose that there exist canonical responses $Y_A, Y_B, Y_C,$ and $Y_D$ taking values in supports $S_A, S_B, S_C,$ and $S_D$ respectively, such that $Y_D = f(Y_A, Y_B, Y_C)$. Suppose further that there exist $0 \leq \alpha, \beta \leq 1$ and sets $S_A \subseteq S_A, S_B \subseteq S_B,$ and $S_{C,D} \subseteq S_C \times S_D$ such that (i)

$$P[(Y_C, Y_D) \in S_{C,D}] > 0,$$

$$P[Y_A \in S_A, (Y_C, Y_D) \in S_{C,D}] = \alpha,$$

$$P[Y_B \in S_B, (Y_C, Y_D) \in S_{C,D}] = \beta;$$

and (ii) $P[Y_A \in S_A, Y_B \in S_B, (Y_C, Y_D) \in S_{C,D}] \neq \alpha \beta$. Then $Y_A \not\perp Y_B \mid (Y_C, Y_D)$.

A straightforward way to ensure conditions (i) and (ii) is to choose $S_A, S_B,$ and $S_{C,D}$ such that $\alpha \beta > 0$, and $Y_A \in S_A$ and $Y_B \in S_B$ imply $(Y_C, Y_D) \notin S_{C,D}$.

For the successors only case ($C = \emptyset$), Theorem 8.3 gives general conditions under which d.3 holds, involving $f$ and the distributions of $Y_A, Y_B,$ and $Y_D$. To illustrate the usefulness of the general result in this case, our next result provides simple informative primitive conditions for $d.3$.

To state this, we introduce two convenient definitions. First, we specify what we mean by saying that predecessors $(Y_1, Y_2)$ are jointly continuously distributed at a point $(y_1^*, y_2^*)$. For simplicity, we let $Y_1$ and $Y_2$ be scalar. For given $\epsilon > 0$, $y_1^* \in \mathbb{R}$, and $y_2^* \in \mathbb{R}$, define neighborhoods $\mathcal{N}_1(\epsilon) := [y_1^* - \epsilon, y_1^* + \epsilon], \mathcal{N}_2(\epsilon) := [y_2^* - \epsilon, y_2^* + \epsilon]$, and $\mathcal{N}(\epsilon) := \mathcal{N}_1(\epsilon) \times \mathcal{N}_2(\epsilon)$.

Definition 8.1 We say $Y_1$ and $Y_2$ are jointly continuously distributed at $(y_1^*, y_2^*)$ if there exists $\epsilon > 0$ such that if $A \subset \mathcal{N}(\epsilon)$ is Borel measurable, then $\lambda(A) > 0$ implies $P[(Y_1, Y_2) \in A] > 0$, where $\lambda$ denotes Lebesgue measure on $\mathbb{R}^2$.

Assuming that $Y_1$ and $Y_2$ are jointly continuously distributed ensures that both $Y_1$ and $Y_2$ exhibit non-trivial random variation and that neither completely determines the other.

Next, we state a mild restriction on the response function.

Definition 8.2 Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be such that there exist $y_1^* \in \mathbb{R}, y_2^* \in \mathbb{R},$ and $\epsilon > 0$ such that for all $y_1$ in $\mathcal{N}_1(\epsilon)$, $f(y_1, \cdot)$ is strictly monotone on $\mathcal{N}_2(\epsilon)$ and for all $y_2$ in $\mathcal{N}_2(\epsilon)$, $f(\cdot, y_2)$ is strictly monotone on $\mathcal{N}_1(\epsilon)$. Then $f$ is locally strictly monotone at $(y_1^*, y_2^*)$.

As special cases, locally strictly monotone functions can be locally strictly increasing or decreasing. The definition also covers mixed cases where, e.g., for all $y_1$ in $\mathcal{N}_1(\epsilon)$, $f(y_1, \cdot)$ is
strictly decreasing on $\mathcal{N}_2(\epsilon)$ and for all $y_2$ in $\mathcal{N}_2(\epsilon)$, $f(\cdot, y_2)$ is strictly increasing on $\mathcal{N}_1(\epsilon)$. Local strict monotonicity is a mild restriction, sufficient to ensure that $\mathcal{X}_1$ and $\mathcal{X}_2$ both cause $\mathcal{X}_3$ with canonical response $Y_3 = f(Y_1, Y_2)$.

**Corollary 8.4** Let $\mathcal{S}$ be as in Theorem 8.3, with $A = \{1\}, B = \{2\}, C = \emptyset$, and $D = \{3\}$. Suppose that $Y_3 = r_3(Y_1, Y_2)$. Suppose further that for some $(y_1^*, y_2^*)$, $Y_1$ and $Y_2$ are jointly continuously distributed at $(y_1^*, y_2^*)$ and that $r_3: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is both continuous and locally strictly monotone at $(y_1^*, y_2^*)$. Then $Y_1 \not\perp Y_2 \mid Y_3$.

This delivers d.3 without imposing the local Markov property. For brevity, we leave further investigation of d.3 and its conditional extension to future work.

## 9 Conclusion

We study the connections between conditional independence and causal relations within the settable systems extension of the Pearl Causal Model. We address concerns raised by Dawid (2002, 2010a, 2010b) to demonstrate how the settable systems framework permits a clear separation between causal and probabilistic concepts and that, while helpful, graphical representations are not essential for the study of these concepts and their interrelations. We provide formal function-based definitions of direct and indirect causality as well as notions of causality via a set of variables and exclusive of a set of variables. These definitions complement and extend the definitions provided in Robins and Greenland (1992), SGS, Pearl (2000, 2001), Robins (2003), Avin, Shpitser, and Pearl (2005), Didelez, Dawid, and Geneletti (2006), and Geneletti (2007). We state and prove the conditional Reichenbach principle of common cause, formally establishing the classical Reichenbach principle as a corollary. We introduce concepts of conditional causal and stochastic isolation to distinguish between situations in which causal restrictions among settable variables ensure that their responses are conditionally independent for any probability measure and those where conditional independence is due to a particular choice of probability measure. These notions yield necessary and sufficient conditions for conditional dependence among specified random vectors in settable systems. We relate our results to the (Markovian) PCM and concepts of $d$-separation and $D$-separation, and we provide conditions under which the causal intuitions these notions support fail or hold.

Taken together, our results show that recursive settable systems provide an effective formal framework for studying the relations between functionally defined causal relations
and conditional independence, and that background variables, the Markov properties, enhanced bases, chance and deterministic nodes, and the assumption of faithfulness or stability are not fundamental to establishing these connections. Nevertheless, we demonstrate that these notions may be helpful for understanding conditional independence relations in certain restricted settable systems.

We focus attention here primarily on recursive systems. An interesting direction for further research is to extend our concepts and results to non-recursive systems (see, e.g., Lauritzen and Richardson, 2002; Wermuth and Cox, 2004; WC). Our framework also constitutes an appropriate foundation for studying the identification and estimation of direct, indirect, and “path-specific” causal effects (See Avin, Shpitser, and Pearl, 2005; Didelez, Dawid, and Geneletti, 2006; Geneletti, 2007).

The results of this paper have direct relevance for empirical research by, among other things, providing foundations enabling researchers to identify, justify, and test the validity of covariates in treatment effect estimation, of instruments in instrumental variables estimation, and of predictors in forecasting models. Specifically, White and Lu (2010) study the choice of covariates for treatment effect estimation using the present framework and results. Chalak and White (2007, 2010) use these to study the choice of instruments in extended instrumental variables estimation. White (2006) and White and Chalak (2010) use this framework to analyze the choice of predictors and to provide tests of unconfoundedness. These studies only represent a start on the many evident opportunities; we intend to pursue these in future research.
10 Mathematical Appendix

Proof of Proposition 6.1 We prove the contrapositive. We have:

\[ r_j(z^{0:b_1}(i), z^*_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A^*)] \]

\[ -r_j(z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A))] \]

\[ = r_j(z^{0:b_1}(i), z^*_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A^*)], \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A^*)]] \]

\[ -r_j(z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A))] \]

Suppose \( \mathcal{X}_i \not\xRightarrow{\mathcal{D}} \mathcal{X}_j \). Then by Definition 6.1, for all admissible interventions with the following corresponding responses for \( \mathcal{X}_j \), we have

\[ r_j(z^{0:b_1}(i), z^*_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A^*)]] \]

\[ -r_j(z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A))] = 0. \]

Also, suppose \( \mathcal{X}_i \not\xRightarrow{S} \mathcal{X}_j \). Then by Definition 6.3, for all admissible interventions with the following corresponding responses for \( \mathcal{X}_j \), we have

\[ r_j(z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A^*)]] \]

\[ -r_j(z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, y_A, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A))] = 0. \]

Since the space of jointly admissible setting values of the form

\[ (z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A)), \]

\[ r_{A:j}[z^{0:b_1}(i), z_i, z_{:A}, z_A, y_A^*, r_{\mathcal{P}}(z_0, z^{1:b_1}(i), z_i, z_{:A}, y_A))] \]

49
includes the space of jointly admissible setting values of the form

\[
(z_{[0:b_1]}(i), z_i; z_{i:A}, z_{\Delta}, y_A; r_{\mathcal{I}}(z_0, z_{[1:b_1]}(i), z_i, z_{i:A}, y_A),
\]

\[
r_{A:j}[z_{[0:b_1]}(i), z_i; z_{i:A}, z_{\Delta}, y_A; r_{\mathcal{I}}(z_0, z_{[1:b_1]}(i), z_i, z_{i:A}, y_A))],
\]

it follows that for all admissible interventions with the following corresponding responses for \( X_j \), we have

\[
r_j(z_{[0:b_1]}(i), z_{\Delta}, z_{i:A}, y_A; r_{\mathcal{I}}(z_0, z_{[1:b_1]}(i), z_i, z_{i:A}, y_A),
\]

\[
r_{A:j}[z_{[0:b_1]}(i), z_i; z_{i:A}, z_{\Delta}, y_A; r_{\mathcal{I}}(z_0, z_{[1:b_1]}(i), z_i, z_{i:A}, y_A))]
\]

\[-r_j(z_{[0:b_1]}(i), z_i; z_{i:A}, z_{\Delta}, y_A; r_{\mathcal{I}}(z_0, z_{[1:b_1]}(i), z_i, z_{i:A}, y_A),
\]

\[r_{A:j}[z_{[0:b_1]}(i), z_i; z_{i:A}, z_{\Delta}, y_A; r_{\mathcal{I}}(z_0, z_{[1:b_1]}(i), z_i, z_{i:A}, y_A))]) = 0,
\]

that is, \( X_i \not\models_S X_j \). This verifies the contrapositive, so the claimed result follows. ■

**Proof of Corollary 6.2** Apply Proposition 6.1 with \( A = \text{ind}(\mathcal{I}_{ij}) \). ■

**Proof of Proposition 6.3** The proof is analogous to that for Proposition 6.1 and is omitted. ■

We next collect together useful basic results on (indirect) causality via or exclusive of \( X_A \) for the special cases \( A = \emptyset \) or \( A = \text{ind}(\mathcal{I}_{ij}) \).

**Proposition 10.1** Let \( S, i, \) and \( j \) be as Definition 6.3. Let \( A = \emptyset \) and \( B = \text{ind}(\mathcal{I}_{ij}) \). Then

(a) \( X_i \not\models_S X_j \);

(b) \( X_i \not\models_S X_j \);

(c) \( X_i \not\equiv_S X_j \) if and only if \( X_i \not\models_S X_j \);

(d) \( X_i \not\models_S X_j \) if and only if \( X_i \not\models_S X_j \);

(e) \( X_i \not\equiv_S X_j \) if and only if \( X_i \not\equiv_S X_j \);

(f) \( X_i \not\equiv_S X_j \) if and only if \( X_i \not\equiv_S X_j \).

It follows from (e) and (f) in Proposition 10.1 that \( \emptyset \)–causality and \( \sim A \)–causality with \( A = \text{ind}(\mathcal{I}_{ij}) \) are equivalent to direct causality in recursive systems.

**Proof of Proposition 10.1**

(a) We have \( \text{ind}(\mathcal{I}_{ij}) = A \), and thus \( r_j(z_{[0:b_2-1]}(i), z_i, z_{\Delta}) \). It follows from \( I[A] \)

Definition 6.3 that \( X_i \not\models_S X_j \) since \( r_j(z_{[0:b_1]}(i), z_i, z_{\Delta}) - r_j(z_{[0:b_1]}(i), z_i, z_{\Delta}) = 0 \) for all function arguments.

50
(b) We have $B = \text{ind}(\mathcal{I}_{i;j})$, and thus $r_j(\bar{z}_{[0:b_2-1]}) = r_j(\bar{z}_{[0:b_1]}(i), z_i, z_B)$. It follows from Definition 6.4 that $\mathcal{X}_i \not\xRightarrow{[\sim B]} \mathcal{X}_j$ since $r_j(\bar{z}_{[0:b_1]}(i), z_i, z_B) - r_j(\bar{z}_{[0:b_1]}(i), z_i, z_B) = 0$ for all function arguments.

(c) Definition 6.4 gives that $\mathcal{X}_i \xRightarrow{[\sim A]} \mathcal{X}_j$ if there exists an admissible intervention $(\bar{z}_{[0:b_1]}(i), z_i, y_A) \rightarrow (\bar{z}_{[0:b_1]}(i), z_i, y_A^*)$ such that

$$r_j(\bar{z}_{[0:b_1]}(i), z_i, y_A^*) - r_j(\bar{z}_{[0:b_1]}(i), z_i, y_A) \neq 0.$$ 

Also, Definition 6.3 gives that $\mathcal{X}_i \xRightarrow{[B]} \mathcal{X}_j$ if there exists an admissible intervention $(\bar{z}_{[0:b_1]}(i), z_i, y_B) \rightarrow (\bar{z}_{[0:b_1]}(i), z_i, y_B^*)$ such that

$$r_j(\bar{z}_{[0:b_1]}(i), z_i, y_B^*) - r_j(\bar{z}_{[0:b_1]}(i), z_i, y_B) \neq 0.$$ 

But we have $A = \text{ind}(\mathcal{I}_{i;j}) = B$. The claim is verified, as the two definitions coincide.

(d) Definition 6.6 gives that $\mathcal{X}_i \xRightarrow{[A]} \mathcal{X}_j$ if there exists an admissible intervention $(\bar{z}_{[0:b_1]}(i), z_i, y_A) \rightarrow (\bar{z}_{[0:b_1]}(i), z_i^*, y_A^*)$ such that

$$r_j(\bar{z}_{[0:b_1]}(i), z_i^*, y_A^*) - r_j(\bar{z}_{[0:b_1]}(i), z_i, y_A) \neq 0.$$ 

Also, Definition 6.5 gives that $\mathcal{X}_i \xRightarrow{[B]} \mathcal{X}_j$ if there exists an admissible intervention $(\bar{z}_{[0:b_1]}(i), z_i, y_B) \rightarrow (\bar{z}_{[0:b_1]}(i), z_i^*, y_B^*)$ such that

$$r_j(\bar{z}_{[0:b_1]}(i), z_i^*, y_B^*) - r_j(\bar{z}_{[0:b_1]}(i), z_i, y_B) \neq 0.$$ 

But we have $A = \text{ind}(\mathcal{I}_{i;j}) = B$. The claim is verified, as the two definitions coincide.

(e) Definition 6.6 gives that $\mathcal{X}_i \xRightarrow{[A]} \mathcal{X}_j$ if there exists an admissible intervention $(\bar{z}_{[0:b_1]}(i), z_i, z_A) \rightarrow (\bar{z}_{[0:b_1]}(i), z_i^*, z_A^*)$ such that

$$r_j(\bar{z}_{[0:b_1]}(i), z_i^*, z_A^*) - r_j(\bar{z}_{[0:b_1]}(i), z_i, z_A) \neq 0.$$ 

Also, Definition 6.5 gives that $\mathcal{X}_i \xRightarrow{[B]} \mathcal{X}_j$ if there exists an admissible intervention $(\bar{z}_{[0:b_1]}(i), z_i, z_B) \rightarrow (\bar{z}_{[0:b_1]}(i), z_i^*, z_B^*)$ such that

$$r_j(\bar{z}_{[0:b_1]}(i), z_i^*, z_B^*) - r_j(\bar{z}_{[0:b_1]}(i), z_i, z_B) \neq 0.$$ 

But we have $A = \text{ind}(\mathcal{I}_{i;j}) = B$. The claim is verified, as the two definitions coincide.

(f) Definition 6.5 gives that $\mathcal{X}_i \xRightarrow{[A]} \mathcal{X}_j$ if there exists an admissible intervention $(\bar{z}_{[0:b_1]}(i), z_i, z_A) \rightarrow (\bar{z}_{[0:b_1]}(i), z_i^*, z_A^*)$ such that

$$r_j(\bar{z}_{[0:b_1]}(i), z_i^*, z_A^*) - r_j(\bar{z}_{[0:b_1]}(i), z_i, z_A) \neq 0.$$ 

51
Also, Definition 6.1 gives that $X_i \xrightarrow{D} S X_j$ if there exists an admissible intervention 

$$(z_{[0:b_1]}(i), z_i, z_{[b_1+1:b_2-1]}) \rightarrow (z_{[0:b_1]}(i), z_i^*, z_{[b_1+1:b_2-1]})$$

such that 

$$r_j(z_{[0:b_1]}(i), z_i^*, z_{[b_1+1:b_2-1]}) - r_j(z_{[0:b_1]}(i), z_i, z_{[b_1+1:b_2-1]}) \neq 0.$$ 

But we have $A = \text{ind}(I_{i;j}) = \Pi_{[b_1+1:b_2-1]}$. The claim is verified, as the two definitions coincide. 

We next state a lemma that plays a key role in formalizing the conditional Reichenbach principle of common cause. For $i \in \Pi_{b_1}$, $j \in \Pi_{b_2}$, $b_1 < b_2$, and $A \subseteq \text{ind}(I_{i;j})$, we let the elements of $y_{A;j} = r_{A;j}(z_{[0:b_1]}(i), z_i, y_{i:A}, y_{A}, y_{\overline{A}})$ obtain by recursive substitution.

**Lemma 10.2** Let $S$ be recursive, $j \in \Pi_b$, and $A$ a subset of $\text{ind}(I_{0;j})$. Let $X_j$ have canonical response $Y_j^c$. If $X_0 \sim^A S X_j$, then there exists a measurable function $\tilde{r}_j$ such that for all admissible setting values

$$r_j(z_0, y_{0:A}, y_{A}, z_A, r_{A}(z_0, y_{0:A}, z_A), r_{A;j}(z_0, y_{0:A}, y_{A}, z_A, r_{A}(z_0, y_{0:A}, z_A))) = \tilde{r}_j(z_A),$$

and in particular

$$y_j^c := r_j(z_0, y_{0:A}, y_{A}, y_{\overline{A}}, y_{A;j}) = \tilde{r}_j(y_A^c).$$

(Recall that here $i = 0$, so that $z_0, y_{0:A}, y_A, y_{\overline{A}}, y_{A;j}$ are canonical setting/response values.)

Thus, if $X_0 \sim^A S X_j$ then, provided it exists, we can express a canonical response $Y_j^c$ as a function of canonical responses $Y_A^c$.

**Proof of Lemma 10.2** Denote by $S^*$ the space of jointly admissible settings for $\langle X_0, P_{0;j}, X_A, X_{\overline{A}}, S_{0;j} \rangle$ of the form $(z_0^*, y_{0:A}^*, y_{A}^*, z_A, r_{A}(z_0^*, y_{0:A}^*, z_A), r_{A;j}(z_0^*, y_{0:A}^*, y_{A}^*, z_A, r_{A}(z_0^*, y_{0:A}^*, z_A)))$. Since $Y_j^c$ exists, $S^*$ is not empty.

First, suppose that $S^*$ is a singleton. Then there does not exist an admissible intervention to $\langle X_0, P_{0;j}, X_A, X_{\overline{A}}, S_{0;j} \rangle$ of the specified form and thus $X_0 \sim^A S X_j$. It follows trivially that there exists a measurable function $\tilde{r}_j$ such that

$$r_j(z_0, y_{0:A}, y_{A}, z_A, r_{A}(z_0, y_{0:A}, z_A), r_{A;j}(z_0, y_{0:A}, y_{A}, z_A, r_{A}(z_0, y_{0:A}, z_A))) = \tilde{r}_j(z_A),$$

and in particular

$$y_j^c = r_j(z_0, y_{0:A}, y_A, y_{\overline{A}}, y_{A;j}) = \tilde{r}_j(y_A^c).$$
Second, suppose that $S^*$ is a multi-element set and that $X_0 \sim^A S X_j$. Then by Definition 6.6 for all admissible interventions to $(X_0, P_{0,j}^A, X_A, X_j, S_{0,j}^A)$ with the following corresponding responses for $X_j$ we have

$$r_j(z_0^*, y_{0:A}^*, y_A^*, z_A^*, r_A(z_0^*, y_{0:A}^*, z_A^*, A), r_{A:j}[z_0^*, y_{0:A}^*, y_A^*, z_A^*, r_A(z_0^*, y_{0:A}^*, z_A)]) = 0.$$ 

Therefore there exists a measurable function $z_A \rightarrow \tilde{r}_j(z_A)$ such that for all elements of $S^*$

$$r_j(z_0, y_{0:A}, y_A, z_A, r_A(z_0, y_{0:A}, z_A), r_{A:j}[z_0, y_{0:A}, y_A, z_A, r_A(z_0, y_{0:A}, z_A)]) = \tilde{r}_j(z_A).$$

In particular, for all $(z_0, y_{0:A}, y_A, y_j, y_{A:j}) \in S^*$ we have

$$r_j(z_0, y_{0:A}, y_A, y_j, y_{A:j}) = \tilde{r}_j(y_A) = \tilde{r}_j(y_A^{c}).$$

A special case of Lemma 10.2 occurs when $A = \emptyset$.

**Corollary 10.3** Let $S$ and $Y_j$ be as in Lemma 10.2. If $X_0 \approx S X_j$ then $Y_j^c$ is constant.

**Proof of Corollary 10.3** Let $A = \emptyset$. Proposition 10.1(d) gives that $X_0 \sim^A S X_j$ if and only if $X_0 \approx_S X_j$. Since $\tilde{r}_j(z_A)$ must be constant, the result follows from Lemma 10.2. $\blacksquare$

**Proof of Proposition 7.1** Apply Theorem 10.4 (see below) with $A = \{i\}$ and $B = \{j\}$. $\blacksquare$

**Proof of Corollary 7.2** Apply Proposition 7.1 with $A = \emptyset$. The result follows from Corollary 10.3. $\blacksquare$

**Proof of Proposition 7.3** We prove the contrapositive.

(i) Suppose that $i = 0$ and that there does not exist an $(X_0, X_j)$ path that does not contain elements of $X_A$. Let $A_j = A \cap ind(\mathcal{I}_{0:j})$. Denote by $S^*$ the space of jointly admissible settings to $(X_0, P_{0,j}^A, X_A, X_j, S_{0,j}^A)$ of the form

$$(z_0^*, y_{0:A}^*, y_A^*, z_A^*, r_A(z_0^*, y_{0:A}^*, z_A)), r_{A:j}[z_0^*, y_{0:A}^*, y_A^*, z_A, r_A(z_0^*, y_{0:A}^*, z_A)]).$$

Since $Y_j^c$ exists, $S^*$ is not empty.

First, suppose that $S^*$ is a singleton. Then there does not exist an admissible intervention to $(X_0, P_{0,j}^A, X_A, X_j, S_{0,j}^A)$ of the specified form and thus $X_0 \sim^A S X_j$, a contradiction.
Second, suppose that $S^*$ is a multi-element set. By construction, for all admissible interventions to $(\mathcal{X}_0, \mathcal{P}^{A_j}_{ij}, \mathcal{X}_{A_j}, \mathcal{X}_{\overline{A_j}}, S^*_{ij})$ with the following corresponding responses for $\mathcal{X}_j$ we have

$$r_j(z^*_0, y_{0,A_j}^*, y_{\overline{A_j}}^*, z_{A_j}, r_{A_j}(z^*_0, y_{0,A_j}^*, y_{\overline{A_j}}^*, z_{A_j}, r_{A_j}(z^*_0, y_{0,A_j}^*, y_{\overline{A_j}}^*, z_{A_j}))$$

$$-r_j(z_0, y_{0,A_j}, y_{\overline{A_j}}, z_{A_j}, r_{A_j}(z_0, y_{0,A_j}, y_{\overline{A_j}}, z_{A_j}), r_{A_j}(z_0, y_{0,A_j}, y_{\overline{A_j}}, z_{A_j}, r_{A_j}(z_0, y_{0,A_j}, y_{\overline{A_j}}, z_{A_j}))) = 0.$$  

Otherwise, it follows from Definition 6.1 of direct causality that there must exist an $(\mathcal{X}_0, \mathcal{X}_j)$ path that does not contain elements of $\mathcal{X}_{A_j}$ and therefore of $\mathcal{X}_A$ by definition of $A_j$. It follows from Definition 6.6 that $\mathcal{X}_0 \not\Rightarrow S \mathcal{X}_j$, a contradiction.

(ii) Suppose that $j = 0$ and that there does not exist an $(\mathcal{X}_0, \mathcal{X}_i)$ path that does not contain elements of $\mathcal{X}_A$. Then an argument parallel to (i) leads to $\mathcal{X}_0 \not\Rightarrow S \mathcal{X}_i$ (with $A_i := A \cap \text{ind}(\mathcal{T}_{0;i})$), a contradiction.

(iii) Suppose that $i, j \neq 0$ and that there does not exist (a) an $(\mathcal{X}_0, \mathcal{X}_i)$ path that does not contain elements of $\mathcal{X}_A$ or (b) an $(\mathcal{X}_0, \mathcal{X}_j)$ path that does not contain elements of $\mathcal{X}_A$ (or both). Then arguments parallel to (i) or (ii) (or both) imply that $\mathcal{X}_0 \not\Rightarrow S \mathcal{X}_j$ or $\mathcal{X}_0 \not\Rightarrow S \mathcal{X}_i$ (or both), a contradiction. ■

Proof of Corollary 7.6 The result is immediate from Proposition 7.1 and the contrapositive of the definition of conditional stochastic isolation. ■

Conditional Reichenbach for the Vector Case

In applications, we are often interested in conditional independence relations between vectors. For this, we first extend the meaning of the notations $\mathcal{X}_i \Rightarrow_{S} \mathcal{X}_j$ and $\mathcal{X}_i \not\Rightarrow_{S} \mathcal{X}_j$ to accommodate disjoint sets of multiple settable variables appearing on the right and left hand sides. For example, if $A$ and $B$ are non-empty disjoint collections of indexes, we let $\mathcal{X}_A$ be a vector of settable variables whose indexes belong to $A$ and similarly for $\mathcal{X}_B$, and we write $\mathcal{X}_A \Rightarrow_{S} \mathcal{X}_B$ if $\mathcal{X}_i \Rightarrow_{S} \mathcal{X}_j$ for some $i \in A \cap \Pi_a$ and $j \in B \cap \Pi_b$ with $a < b$. Otherwise, we write $\mathcal{X}_A \not\Rightarrow_{S} \mathcal{X}_B$ indicating that $\mathcal{X}_i \not\Rightarrow_{S} \mathcal{X}_j$ for all $i \in A \cap \Pi_a$ and $j \in B \cap \Pi_b$ with $a < b$. Observe that even though $S$ is a recursive system, it is possible to have $\mathcal{X}_A \Rightarrow_{S} \mathcal{X}_B$ and $\mathcal{X}_B \not\Rightarrow_{S} \mathcal{X}_A$.

Similarly, we extend the notations $\Rightarrow_{S}$, $\Rightarrow_{S}^*$, $\not\Rightarrow_{S}$, and $\not\Rightarrow_{S}^*$ and their negations to accommodate disjoint sets of multiple settable variables appearing on the right and left hand sides. To do this requires some further notation. Let $\mathcal{T}_{A,B} = \cup_{i \in A} \cup_{j \in B} \mathcal{T}_{ij} \setminus (\mathcal{X}_A \cup \mathcal{X}_B)$ denote the set of $(\mathcal{X}_A, \mathcal{X}_B)$–intercessors and let $C \subset \text{ind}(\mathcal{T}_{A,B})$. For given $i \in A$ and $j \in B$, let $C_{ij} = C \cap \text{ind}(\mathcal{T}_{ij})$. Then, we say that $\mathcal{X}_A \Rightarrow_{S} \mathcal{X}_B$ if there exists $i \in A \cap \Pi_a$ and
The definitions of conditional causal isolation and conditional \( P \)-stochastic isolation generalize to the vector case in the obvious way. Thus, if \( Y_A^c \perp Y_B^c \mid Y_C^c \) when \( \mathcal{X}_A \) and \( \mathcal{X}_B \) are not causally isolated given \( \mathcal{X}_C \) (that is, condition (a) in Theorem 10.4 below holds), then we say that \( \mathcal{X}_A \) and \( \mathcal{X}_B \) are \( P \)-stochastically isolated given \( \mathcal{X}_C \).

**Theorem 10.4 Conditional Reichenbach Principle of Common Cause (III)** Let \( S \) be recursive. Let \( A \) and \( B \) be non-empty disjoint subsets of \( \Pi \cup \Pi_0 \) and \( C \subset \Pi \setminus (A \cup B) \). Let \( \mathcal{X}_A \), \( \mathcal{X}_B \), and \( \mathcal{X}_C \) be the corresponding vectors of settable variables with canonical responses \( Y_A^c \), \( Y_B^c \), and \( Y_C^c \). For given probability measure \( P \) on \( (\Omega, \mathcal{F}) \), \( Y_A^c \nmid Y_B^c \mid Y_C^c \) if and only if (a) either:

(i) \( 0 \in A \) and \( \mathcal{X}_0 \) causes \( \mathcal{X}_B \) exclusive of \( C_B := C \cap \text{ind}(\mathcal{I}_{\{0\}}; B) \), i.e., \( \mathcal{X}_0 \Rightarrow_{S} \mathcal{X}_B \); or

(ii) \( 0 \in B \) and \( \mathcal{X}_0 \) causes \( \mathcal{X}_A \) exclusive of \( C_A := C \cap \text{ind}(\mathcal{I}_{\{0\}}; A) \), i.e., \( \mathcal{X}_0 \Rightarrow_{S} \mathcal{X}_A \); or

(iii) \( 0 \not\in A \cup B \) and \( \mathcal{X}_0 \sim_{C_B} \mathcal{X}_A \) and \( \mathcal{X}_0 \sim_{C_A} \mathcal{X}_B \);

and (b) \( \mathcal{X}_A \) and \( \mathcal{X}_B \) are not \( P \)-stochastically isolated given \( \mathcal{X}_C \).

**Proof of Theorem 10.4** Let \( P \) be any probability measure. First, we prove that if \( Y_A^c \nmid Y_B^c \mid Y_C^c \), then \( \mathcal{X}_A \) and \( \mathcal{X}_B \) are not causally isolated given \( \mathcal{X}_C \).

(i) Suppose that \( 0 \in A \) and \( \mathcal{X}_0 \sim_{C_B} \mathcal{X}_B \). Then \( \mathcal{X}_0 \Rightarrow_{S} \mathcal{X}_j \) for all \( j \in B \). For given \( j \in B \), let \( \mathcal{X}_{0,j} \) be a vector of settable variables and let \( Y_{C_{0,j}}^c \) denote the corresponding canonical responses. By Lemma 10.2, it follows that \( Y_j^c = \tilde{r}_j(Y_{C_{0,j}}^c) \) for all \( j \in B \). Let \( \mathcal{X}_{C_B} \) be a vector of settable variables and let \( Y_{C_B}^c \) denote the corresponding canonical response; then we have \( Y_B^c = \tilde{r}_B(Y_{C_B}^c) \). Let \( C_B^c := C \setminus C_B \), let \( \mathcal{X}_{C_B}^c \) be a vector of settable variables, and let \( Y_{C_B}^c \) denote the corresponding canonical responses; then \( Y_C^c = (Y_{C_B}^c, Y_{C_B}^c)^c \). Since \( Y_B^c = \tilde{r}_B(Y_{C_B}^c) \), we have that \( (Y_A^c, Y_{C_B}^c)^c \perp Y_B^c \mid Y_{C_B}^c \). We then have that \( Y_A^c \perp Y_B^c \mid Y_{C_B}^c \) (see, for example, Dawid, 1979, section 4; Döhler, 1981, lemma 3; Smith 1989, property 3; and Florens, Mouchart, and Rolin 1990, theorem 2.2.10), that is, \( Y_A^c \perp Y_B^c \mid Y_{C_B}^c \), a contradiction. (Note that when \( C_B = C \) the result is immediate. Also, when \( C_B = \emptyset \), \( Y_B \) is constant and the result is trivial.)

(ii) Suppose \( 0 \in B \), and that \( \mathcal{X}_0 \sim_{C_A} \mathcal{X}_A \). The result is symmetric to (i) yielding that \( Y_A^c \perp Y_B^c \mid Y_{C_B}^c \), a contradiction.

(iii) Suppose that \( 0 \not\in A \cup B \), and that \( \mathcal{X}_0 \sim_{C_A} \mathcal{X}_A \) or \( \mathcal{X}_0 \sim_{C_B} \mathcal{X}_B \). Suppose that \( \mathcal{X}_0 \sim_{C_A} \mathcal{X}_A \); then an argument similar to (i) gives that \( Y_A^c \perp Y_B^c \mid Y_{C_B}^c \), a contradiction.
Alternatively, suppose that \( \mathcal{X}_0 \overset{C}{\not\equiv}_S \mathcal{X}_B \). Then by a parallel argument, we obtain that 
\( Y_A \perp Y_B \mid Y_C \), a contradiction.

That \( \mathcal{X}_A \) and \( \mathcal{X}_B \) are not stochastically isolated given \( \mathcal{X}_C \) follows by the definition of conditional stochastic isolation. The rest of the proof follows from (the contrapositive of) the definition of conditional stochastic isolation. ■

**Proof of Proposition 8.1** Let \( k \in \Pi_1 \) such that \( \mathcal{X}_0 \not\overset{\sim}{\rightarrow}_S \mathcal{X}_i \). By construction we have that \( \mathcal{X}_0 \not\overset{\sim}{\rightarrow}_S \mathcal{X}_i \). Theorem 10.4 gives that \( Y_i \perp Y_A \mid (Y_C, Y_k) \). Further, since elements of \( \mathcal{X}_C \) and \( \mathcal{X}_A \) do not succeed \( \mathcal{X}_i \), there exists a set \( D \subset \Pi_1 \setminus \{k\} \) such that 
\( \mathcal{X}_0 \not\overset{\sim}{\rightarrow}_S (\mathcal{X}_C, \mathcal{X}_A) \). It follows from Lemma 10.2 that there exists a measurable function \( \tilde{r}_{C,A} \) such that \( (y_C, y_A) = \tilde{r}_{C,A}(y_D) \). Since \( \{Y_k : k \in \Pi_1\} \) are jointly independent we have that \( Y_k \perp Y_D \). It follows from Dawid (1979, lemma 4.2(i)) that \( Y_k \perp (Y_C, Y_A) \). Also, Dawid, 1979, section 4 (see also Döhler, 1981, lemma 3; Smith 1989, property 3; and Florens, Mouchart, and Rolin 1990, theorem 2.2.10) gives that \( Y_k \perp Y_A \mid (Y_C, Y_k) \).

Dawid (1979, lemma 4.3) gives that \( Y_i \perp Y_A \mid Y_C \). ■

**Proof of Proposition 8.2** (i) By construction we have that \( \mathcal{X}_0 \not\overset{\sim}{\rightarrow}_S \mathcal{X}_i \). It follows from Theorem 10.4 that \( Y_i \perp Y_A \mid Y_C \). (ii) An argument similar to Proposition 8.1 gives that \( Y_i \perp Y_{A_1} \mid Y_C \). ■

**Proof of Theorem 8.3** Since \( P[(Y_C, Y_D) \in S_{C,D}] > 0 \), it follows that

\[
P[Y_A \in S_A \mid (Y_C, Y_D) \in S_{C,D}] = \frac{P[Y_A \in S_A, (Y_C, Y_D) \in S_{C,D}]}{P[(Y_C, Y_D) \in S_{C,D}]} = \frac{\alpha}{P[(Y_C, Y_D) \in S_{C,D}]} \quad \text{and} \quad P[Y_B \in S_B \mid (Y_C, Y_D) \in S_{C,D}] = \frac{P[Y_B \in S_B, (Y_C, Y_D) \in S_{C,D}]}{P[(Y_C, Y_D) \in S_{C,D}]} = \frac{\beta}{P[(Y_C, Y_D) \in S_{C,D}]},
\]

so

\[
P[Y_A \in S_A \mid (Y_C, Y_D) \in S_{C,D}] \times P[Y_B \in S_B \mid (Y_C, Y_D) \in S_{C,D}] = \frac{\alpha \beta}{P[(Y_C, Y_D) \in S_{C,D}]}.
\]

Now \( P[Y_A \in S_A, Y_B \in S_B, (Y_C, Y_D) \in S_{C,D}] \neq \alpha \beta \) and \( P[(Y_C, Y_D) \in S_{C,D}] > 0 \) imply

\[
P[Y_A \in S_A, Y_B \in S_B \mid (Y_C, Y_D) \in S_{C,D}] = \frac{P[Y_A \in S_A, Y_B \in S_B, (Y_C, Y_D) \in S_{C,D}]}{P[(Y_C, Y_D) \in S_{C,D}]} \neq \frac{\alpha \beta}{P[(Y_C, Y_D) \in S_{C,D}]}.
\]

It follows that

\[
P[Y_A \in S_A \mid (Y_C, Y_D) \in S_{C,D}] \times P[Y_B \in S_B \mid (Y_C, Y_D) \in S_{C,D}] \neq P[Y_A \in S_A, Y_B \in S_B \mid (Y_C, Y_D) \in S_{C,D}],
\]

56
which implies that $Y_A \not\in Y_B \mid (Y_C, Y_D)$.

**Proof of Corollary 8.4** For brevity, we consider only the locally strictly increasing case. The other cases are similar. We choose sets $S_1$, $S_2$, and $S_3$ such that $P[Y_3 \in S_3] > 0$, $P[Y_1 \in S_1, Y_3 \in S_3] = \alpha > 0$, $P[Y_2 \in S_2, Y_3 \in S_3] = \beta > 0$, and $P[Y_1 \in S_1, Y_2 \in S_2, Y_3 \in S_3] = 0$. The result then follows from Theorem 8.3.

For the given $\epsilon > 0$, $y_1^*$, and $y_2^*$, let $N^-(\epsilon) := [y_1^* - \epsilon, y_1^*]$, $N^+(\epsilon) := [y_1^* + \epsilon, y_1^* + \epsilon]$, $N^-(\epsilon) := [y_2^* - \epsilon, y_2^*]$, and $N^+(\epsilon) := [y_2^* + \epsilon, y_2^*]$. We let $S_1 = N^-(\epsilon)$, $S_2 = N^-(\epsilon)$, and $S_3 := [y_3^*, \bar{y}_3]$, where $y_3^* := r_3(y_1^*, y_2^*)$ and

$$\bar{y}_3 := \min\{r_3(y_1^*, y_2^* + \epsilon), r_3(y_1^* + \epsilon, y_2^*)\}.$$

By the local strictly increasing property, $y_3^* < \bar{y}_3$.

We have $P[Y_3 \in S_3] \geq P[Y_1 \in N^+(\epsilon), Y_2 \in N^+(\epsilon), Y_3 \in S_3] > 0$, as $r_3$ is locally strictly increasing at $(y_1^*, y_2^*)$ and the specified event has positive Lebesgue measure. Next, $P[Y_1 \in S_1, Y_3 \in S_3] \geq P[Y_1 \in N^+(\epsilon), Y_2 \in N^+(\epsilon), Y_3 \in S_3] > 0$ for the same reasons. Similarly, $P[Y_2 \in S_2, Y_3 \in S_3] \geq P[Y_1 \in N^+(\epsilon), Y_2 \in N^+(\epsilon), Y_3 \in S_3] > 0$. This verifies (i) of Theorem 8.3.

But $P[Y_1 \in S_1, Y_2 \in S_2, Y_3 \in S_3] = P[Y_1 \in N^-(\epsilon), Y_2 \in N^-(\epsilon), Y_3 \in S_3] = 0$ by the local strictly increasing property. This verifies (ii) of Theorem 8.3, and the proof is complete. ■

**References**


