A Calculus for Causal Relevance

Blai Bonet
Cognitive Systems Laboratory
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA 90024
bonet@cs.ucla.edu

Abstract

We present a sound and complete calculus for causal relevance that uses Pearl's functional causal models as semantics. The calculus consists of axioms and rules of inference for reasoning about causal relevance relationships. We extend the set of known axioms for causal relevance with new axioms and rules of inference. The axioms are then divided into different sets for reasoning about specific subclasses of models. These subclasses make up a new decomposition of the class of causal models. At the end, we show how the calculus for causal relevance can be used in the task of identifying causal structure from non-observational data.

1 Introduction

Causal relevance is concerned with questions of the form: “Given that (variable) $Z$ is fixed, would changing $X$ alter $Y$?” Formal interest in this notion appears in the works of Suppes [8] and Salmon [7] who attempted to give it probabilistic interpretation. This paper pursues a logical approach, and starts with the work by Galles and Pearl [2], which, similar to the work on graphoids [5, 3], is based on a set of axioms and rules of inference that defines a formal deductive system. In this system, causal relevance is expressed by logical formulas and new relevance sentences can be derived from old ones through rules of inference.

Such deductive system requires an interpretation; that is, a function that maps models to the sets of formulas they “satisfy”. The models we focus on are Pearl's functional causal models [6]. Once an interpretation is established, a natural question to ask is whether the system is sound (i.e., if every proven statement is true), and whether the system is complete (i.e., if every true statement is provable). In this paper we answer both questions in affirmative for some classes of models.

To do so, we define a hierarchy of causal models (see Halpern [4]), governed by new axioms and rules of inference. Towards the end of the paper (Section 4), we show how information about causal relevance can be used for identifying the structure of causal models from data.

The rest of the paper is organized as follows. In Section 2, we present Pearl's functional causal models, the formal language of causal relevance, the interpretation of the language, and the classes of causal models. Section 3 presents the axiomatizations, some general results about the calculus as well as the proofs for soundness and completeness. Then, Section 4 shows how this theory can be used to solve identification problems. We conclude with a brief summary.

2 Model Theory

A causal model is a set of variables in which each variable may have influence on others. The variables are divided in two groups. The endogenous variables for which the model provides a description of the mechanisms that influence them, and the exogenous variables that influence only endogenous variables and are taken as given; i.e., the model does not explain them. In this section we provide a formal definition of causal models.

Without loss of generality, we can assume that there is one (big) exogenous variable that takes values in a domain for which there are no restrictions. All the endogenous variables for causal models are taken from a finite set $\mathcal{X}$. This set is assumed to be fixed for the rest of the paper and often we use it without mention. Some authors like to say that $\mathcal{X}$ is the signature of the theory. It is a technical necessity used to restrict the set of models and the formal language for causal relevance.

2.1 Causal Models

A causal model $T$ is a tuple $T = (\mathcal{U}, \{F_Y : Y \in \mathcal{X}\})$ where $\mathcal{U}$ is the domain of the unique exogenous vari-
able and \( \{F_Y : Y \in \mathcal{X}\} \) is a set of equations called \textit{modifiable structural equations}. Each equation is a function 
\[
F_Y : \prod_{X \in \mathcal{X}} \mathcal{D}_X \times \mathcal{U} \rightarrow \mathcal{D}_Y,
\]
where \( \mathcal{D}_X \) is the domain of variable \( X \), that describes how the variable \( Y \) attains its value.

Given a subset \( X \subseteq \mathcal{X} \) of endogenous variables, we define the causal (sub)model \([x]\mathcal{T} \), that results from the \textit{intervention} of setting the value of \( X \) to \( x \), as 
\[
[x]\mathcal{T} = (\mathcal{U}, \{F'_Y : Y \in \mathcal{X})\} \text{ where } F'_Y = F_Y \text{ for } Y \notin X \text{ and } F'_Y = x \mathcal{D}_Y \text{ for } Y \in X.
\]
This intervention is denoted by the operator \( \text{do}(X = x) \) that maps causal models into causal models by \( T \sim [x]\mathcal{T} \) (see [6]). A causal model defines a system of equations once the exogenous variable had been set. In general, the system may have zero, one or multiple solutions. We will only consider causal models \( T \) such that for each subset \( X \subseteq \mathcal{X} \) and value \( x \) for \( X \), the submodel \([x]\mathcal{T} \) has a unique solution for all \( u \in \mathcal{U} \). This class of models was denoted by \( \mathcal{T}_{\text{uniq}} \) in [4] (note that \( \mathcal{T}_{\text{uniq}} \) depends on \( \mathcal{X} \)).

The \textit{potential response} for variable \( Y \) and \( u \in \mathcal{U} \) under intervention \( \text{do}(X = x) \) is defined as the solution \( Y \) in model \([x]\mathcal{T} \) when the exogenous variable has value \( u \) (denoted by the function \( u \sim [x]Y(u) \)). The main difference between causal models and Bayesian networks is that the former incorporate a semantics for all interventions \( \text{do}(X = x) \).

A \textit{counterfactual} is a formula built from atomic terms of the form \([x]Y(u) \) (see Pearl [6]). A logical calculus for reasoning about counterfactuals that is sound and complete with respect to \( \mathcal{T}_{\text{uniq}} \) was given in [4]. One result from the counterfactual calculus that is important to us is the following.

\textbf{Lemma 1 (Composition; [2])} Let \( W, X, Y \) be variables in a causal model \( T \), \( x \) a value for \( X \), and \( u \) a valuation for the exogenous variable. If \([x]W(u) = w \) and \([x]Y(u) = y \), then \([x,w]Y(\cdot) = y \).

Composition can be briefly summarized by the equation 
\[
[x]Y(u) = [x, [x]W(u)]Y(u) \text{ which is valid for all variables } Y, W, \text{ all valuations } x \text{ of } X, \text{ and all } u \in \mathcal{U}.
\]
As a remark, note that composition is not about nested interventions, it is just a simple intervention that uses a \textit{previously recorded} value from another intervention.

\section{2.2 Language}

We talk about causal relevance using a simple logical language built from atomic formulas and the standard boolean connectives. The language, that also depends on \( \mathcal{L} \), is denoted by \( \mathcal{L} \). Each atomic formula, or just atom, of \( \mathcal{L} \) is of the form \( (X \rightarrow Y)[Z] \) where \( X, Y \) and \( Z \) are disjoint subsets of \( \mathcal{X} \), and \( X,Y \) are non-empty. Intuitively, an atom \( (X \rightarrow Y)[Z] \) is read as \textit{variable X has no influence on variable Y once variable Z had been fixed}. The language is inductively defined by (i) every atom is in \( \mathcal{L} \), (ii) if \( \varphi \in \mathcal{L} \), then \( \neg \varphi \in \mathcal{L} \), and (iii) if \( \varphi, \psi \in \mathcal{L} \), then \( (\varphi \& \psi), (\varphi \vee \psi), (\varphi \Rightarrow \psi) \in \mathcal{L} \).

As is common, a \textit{literal} is an atom or the negation of an atom. It is positive if it is an atom, otherwise it is negative. We will use capital Greek letters to denote sets of formulas and lower case Greek letters to denote formulas. If \( \Gamma \) is a set of formulas, the set of positive (resp. negative) literals in \( \Gamma \) is denoted by \( \text{Lit}^+ \Gamma \) (resp. \( \text{Lit}^- \Gamma \)).

\section{2.3 Interpretations}

An interpretation assigns to each model the set of formulas that it satisfies. We begin with the interpretation for atoms and then extend it to all formulas.

\textbf{Definition 1} A causal model \( T \in \mathcal{T}_{\text{uniq}} \) satisfies the atom \((X \rightarrow Y)[Z]\) if and only if 
\[
[x, z, w]Y(\cdot) = [x', z, w]Y(\cdot)
\]
for all \( W \) disjoint of \( X, Y, Z \), and all valuations \( x, x', z, w \) for the variables \( X, Z \) and \( W \).

Note that Eq.(1) is an equality of functions so it must hold for all \( u \in \mathcal{U} \). We write \( T \vDash \varphi \) when \( T \) satisfies formula \( \varphi \), and \( T \vDash \Gamma \) when \( T \) satisfies all formulas in \( \Gamma \). The interpretation is extended to all formulas by (i) \( T \vDash \neg \varphi \) if \( T \not\vDash \varphi \), (ii) \( T \vDash (\varphi \& \psi) \) if \( T \vDash \varphi \) and \( T \vDash \psi \), (iii) \( T \vDash (\varphi \vee \psi) \) if \( T \vDash \varphi \) or \( T \vDash \psi \), and (iv) \( T \vDash (\varphi \Rightarrow \psi) \) if \( T \vDash \varphi \) but \( T \not\vDash \psi \).

In the next section, we use the interpretations to decompose the class \( \mathcal{T}_{\text{uniq}} \) into smaller classes of causal models.

\section{2.4 Classes of Models}

We begin with some definitions. The \textit{semantic graph} associated with model \( T \) and \( u \in \mathcal{U} \) is the directed graph \( G(T, u) \) with vertex set equal to \( \mathcal{X} \) and edge set \( E \) such that \( (X, Y) \in E \) if and only if \( X \) is a non-trivial argument of \( F_Y(\cdot; u) \). The semantic graph \( G(T) \) associated with \( T \) is the union of all graphs \( G(T, u) \). Given two graphs \( G, G' \), we write \( G \preceq G' \) when \( G \) is an edge subgraph of \( G' \).

\textbf{Definition 2} A causal model \( T \) is \textit{strong recursive} if \( G(T) \) is a \textit{DAG}. It is \textit{recursive} if \( G(T, u) \) is a \textit{DAG} for all \( u \in \mathcal{U} \).

Let \( \mathcal{T}_{\text{rec}} \) and \( \mathcal{T}_{\text{rec}} \) be the subclasses of models from \( \mathcal{T}_{\text{uniq}} \) that are strong recursive and recursive respectively. By definition, \( \mathcal{T}_{\text{rec}} \subseteq \mathcal{T}_{\text{rec}} \subseteq \mathcal{T}_{\text{uniq}} \).
We say that two models \( T, T' \) are \( L \)-equivalent, written \( T \equiv T' \), if they can not be distinguished by the language; i.e., \( \forall \varphi[T \models \varphi \iff T' \models \varphi] \). We say that a class \( \mathcal{T} \) is \( L \)-\textit{contained} into another class \( \mathcal{T}' \), written \( \mathcal{T} \subseteq \mathcal{T}' \), if \( \forall (T \in \mathcal{T}) \exists (T' \in \mathcal{T'})[T \equiv T'] \). The following diagram shows the partial order induced by \( \subseteq \) among the classes of models (modulo \( \equiv \)). The labels at the arrows indicate whether the relation is strict or not.

\[
\begin{array}{ccc}
\mathcal{T}_{\text{rec}} & \prec & \mathcal{T}_{\text{rec}} \\
\mathcal{T}_{\text{rec}} & \prec & \mathcal{T}_{\text{uniq}}
\end{array}
\]

(2)

Not all the information in this diagram is obvious. The unstrict relations follow from the containments above noted. The fact that \( \mathcal{T}_{\text{rec}} \prec \mathcal{T}_{\text{rec}} \) is seen with the model \( T \) with \( X = \{X_1, X_2\}, U = \{u, u'\} \) and equations:

\[
\begin{align*}
X_1(u) &= X_2(u), \\
X_1(u') &= 0,
\end{align*}
\]

\[
X_2(u) = 0,
\]

\[
X_2(u') = X_1(u');
\]

i.e., with graph \( X_1 \rightleftharpoons X_2 \). The other strict relation is more difficult to see; we will prove it later.

From now on, we will only consider the equivalence classes generated by \( \equiv \). To simplify notation, however, we write \( T \) and \( \mathcal{T} \) instead of \([T]_{\equiv}\) and \([\mathcal{T}]_{\equiv}\) respectively.

3 Axiomatizations

An axiomatic system \( AX \) is a set of formulas (or schemata) from the language plus a set of rules of inference. Each rule of inference is a \textit{license} that allows us to derive new formulas from previous ones. A \textit{proof} for formula \( \varphi \) from a set \( \Gamma \) is a finite sequence of formulas \( \varphi_1, \varphi_2, \ldots, \varphi_n = \varphi \) such that each \( \varphi_i \) is either an axiom, a formula in \( \Gamma \), or a formula derived using a rule of inference from previous ones. We write \( \Gamma \vdash_{AX} \varphi \) if \( \varphi \) can be proved from \( \Gamma \) in the AX system. An axiomatic system AX is \textit{sound} with respect to a class of models \( \mathcal{T} \) if \( \Gamma \vdash_{AX} \varphi \) implies \( \Gamma \vdash T \varphi; \) i.e., everything that can be proved from \( \Gamma \) is valid in all models that satisfy \( \Gamma \). AX is \textit{complete} with respect to \( \mathcal{T} \) if \( \varphi \) implies \( \Gamma \vdash_{AX} \varphi \); i.e., everything that is valid in models that satisfy \( \Gamma \) can be proved from \( \Gamma \).

Axioms A2–A4, A6 and A7 appeared in Pearl [6]. The other three are new.

Let \( \Gamma \subseteq L \) be a set of formulas and fix an axiomatic system AX. The following definitions are standard. \( \Gamma \) is a \textit{closed set} if \( \forall \varphi[\Gamma \vdash_{AX} \varphi \Rightarrow \varphi \in \Gamma] \). The \textit{theory} associated with \( \Gamma \), denoted as \( \text{Th}_{AX} \Gamma \), is the smallest closed set containing \( \Gamma \). \( \Gamma \) is a \textit{consistent} set if \( \forall \varphi[\varphi \in \text{Th}_{AX} \Gamma \Rightarrow \neg \varphi \notin \text{Th}_{AX} \Gamma] \). A set \( \Gamma \supseteq \Gamma \) is an \textit{extension} of \( \Gamma \) if it is consistent and maximal. The theory associated with a model \( T \) is the set of all formulas valid in \( T \); i.e. \( \text{Th} T = \{ \varphi \in L : T \models \varphi \} \).

The AXuniq axiomatic system is defined as the Axioms A1–A9 plus Modus Ponens (MP) as the unique rule of inference. Often we say that \( \Gamma \) is AXuniq-consistent as an abbreviation for \( \Gamma \) is consistent with respect to AXuniq. Similarly for extensions and for other axiomatic systems.

Theorem 2 AXuniq is sound with respect to \( \mathcal{T}_{\text{uniq}} \).

\[
\text{Proof:} \quad \text{The soundness of A1 is obvious. Fix } u \in U. \text{ For clarity, we won't write the dependency in } u. \text{ The soundness of A2, A3, A6 and A7 was proved in [2]. Axiom A4 is in [6] but its proof has not been published. Therefore, we give the proofs only for A4, A5, A8 and A9. In what follows, lower case Roman letters refer to values for the corresponding capital letters; e.g., } x, x' \text{ will be valuations for } X. \text{ In each case, the suppositions are denoted from first to last by S1}, S2, ...\n\]

A4. Suppose \( (X \rightarrow YW|Z) \) and \( (X \rightarrow Y|ZW) \), and let \( V \) be disjoint from \( WXYZ \). Then,

\[
[x, z, v]Y = [x, z, v, [x, z, v]W]Y \quad \text{ (comp.)}
\]
\[
= [x', z, v, [x', z, v]W]Y \quad \text{ (by S2)}
\]
\[
= [x', z, v]WY \quad \text{ (by S1)}
\]
\[
= [x', z, v]Y. \quad \text{ (comp.)}
\]
Thus, \((X \rightarrow Y\mid Z)\).

A5. Suppose \((X \rightarrow YW\mid Z)\) and \((W \rightarrow Y\mid ZX)\), and let \(V\) be disjoint from \(XYWZ\). Then,

\[
[x, z, v, t]Y = [x, z, v, [x, z, v]W]Y \quad \text{(by S2)}
\]

\[
= [x, z, v]Y \quad \text{(comp.)}
\]

\[
= [x', z, v]Y \quad \text{(by S1)}
\]

\[
= [x', z, v, [x', z, v]W]Y \quad \text{(by S2)}
\]

Thus, \((X \rightarrow Y\mid Z)\).

A8. Suppose \((X \rightarrow Y\mid ZV)\), \((X \rightarrow Y\mid ZU)\), \((U \rightarrow V\mid ZXW)\) and \((V \rightarrow U\mid ZXW)\), and let \(T\) be disjoint from \(WXYUV\). Then,

\[
[x, w, z, t]Y
= [x, w, z, t, [x, w, z, t]U]Y \quad \text{(comp.)}
\]

\[
= [x', w, z, t]Y \quad \text{(by S2)}
\]

\[
= [x', w, z, t, [x', w, z, t]U]Y \quad \text{(by S3)}
\]

\[
= [x', w, z, t, [x', w, z, t]U, [x', w, z, t]V]Y \quad \text{(by S4)}
\]

\[
= [x', w, z, t, [x', w, z, t]U, [x', w, z, t]V]Y \quad \text{(by S5)}
\]

\[
= [x', w, z, t, [x', w, z, t]V]Y \quad \text{(by S6)}
\]

\[
= [x', w, z, t]Y. \quad \text{(comp.)}
\]

Thus, \((X \rightarrow Y\mid ZW)\).

A9. Suppose \((X \rightarrow YW\mid Z)\), \((W \rightarrow Y\mid ZX)\) and \((X \rightarrow YV\mid ZW)\). They imply, by A5, A4 and A7, \((X \rightarrow Y\mid Z)\) and \((X \rightarrow YV\mid ZW)\). Therefore, we only need to prove the result for \(V\) and when the context \(T\) is disjoint from \(VXYZ\). Then,

\[
[x, z, t]V
= [x, z, t, [x, z, t]W]V \quad \text{(comp.)}
\]

\[
= [x, z, t, [x, z, t]W, [x, z, t]W]V \quad \text{(by S1)}
\]

\[
= [x', z, t, [x', z, t]W]V \quad \text{(by S2)}
\]

\[
= [x', z, t, [x', z, t]W, [x', z, t]W]V \quad \text{(by S3)}
\]

\[
= [x', z, t]V. \quad \text{(comp.)}
\]

Thus, \((X \rightarrow YV\mid Z)\).

Some Theorems of \(AX_{\text{uniq}}\) are:

1. **Extended Left Intersection:**

\[
(TX \rightarrow Y\mid ZW) \&(TW \rightarrow Y\mid ZX) \Rightarrow (TXW \rightarrow Y\mid Z),
\]

2. **Extended Right Intersection:**

\[
(X \rightarrow TY\mid ZW) \&(X \rightarrow TW\mid ZX) \Rightarrow (X \rightarrow TYW\mid Z).
\]

Let \(\Gamma\) be a maximal consistent set of formulas with respect to \(AX_{\text{uniq}}\). A set \(P\) of variables is called the parent set of \(Y\), also denoted as \(Pa(Y)\), if \(((PY)^e \rightarrow Y\mid P) \in \Gamma\) and \(P\) is \(\subseteq\)-minimal. If such set \(P\) does not exist, then \(Pa(Y)\) is defined as \(Y^c\). The syntactic graph associated with a maximal consistent set \(\Gamma\) is the directed graph \(G(\Gamma)\) with vertex set equal to \(X\) and edge set \(E\) such that \((X, Y) \in E\) if and only if \(X \in Pa(Y)\). From now on, if the set \(\Gamma\) is clear from context, we do not write the suffix "in \(\Gamma\)".

**Theorem 3** Let \(\Gamma\) be a maximal \(AX_{\text{uniq}}\)-consistent set and \(T\) a causal model with exogenous domain \(U\). Then,

(i) \(Pa(Y)\) is unique (i.e., \(G(\Gamma)\) is well defined),

(ii) \(X \in Pa(Y)\) iff \((X \rightarrow Y\mid W)\) for all \(W\),

(iii) \(G(T) = G(Th T)\),

(iv) if \((X \rightarrow Y\mid Z)\), then there is a (directed) path \(X \rightarrow Y\) not intersected by \(Z\) in \(G(\Gamma)\).

*Proof:* (i). It is enough to show that the collection of sets that "separates" \(Y\) from the rest is closed under intersections. Let \(W_1, P_1\) and \(W_2, P_2\) be sets of variables such that \(W_1 = (P_1Y)^e\), \(W_2 = (P_2Y)^e\), \((W_1 \rightarrow Y\mid P_1)\) and \((W_2 \rightarrow Y\mid P_2)\). Let \(P = P_1 \cap P_2\). Then, by \(P_1 \subseteq W_2, P_2 \subseteq W_1\) and \(Ext\) Left Intersection, \((W \rightarrow Y\mid P)\) where \(W = (YP)^e\).

(ii). Let \(P = Pa(Y)\) \(\setminus\) \(X, W = (XY)^e\). Then, 

\[
\neg(WX \rightarrow Y\mid P)\quad \text{and, \ } A_6, \neg(X \rightarrow Y\mid WP).
\]

If \(X \notin Pa(Y)\), then by \(A_2\) and definition, \(X \notin Pa(Y)\).

(iii). Need to prove that an edge \((X, Y) \in G(T)\) if and only if \((X, Y) \in G(Th T)\). Observe that \(Th T\) is maximal consistent. Let \(W = (XY)^e\). Then,

\[
(X, Y) \in G(T) \iff \exists u((X, Y) \in G(T, u))
\]

\[
\iff \exists u((X, Y) \in G(T, u))
\iff \exists u(X \text{ is non-trivial arg. of } F_Y(\cdot; u))
\iff \exists u(T(u) = -(X \rightarrow Y\mid W)]
\iff \neg(X \rightarrow Y\mid W) \in Th T
\iff X \in Pa(Y).
\]

(iv). By reverse induction in \#Z (cardinality of \(Z\)). Let \(Z = (XY)^e\) such that \(\neg(X \rightarrow Y\mid Z)\). Then, \(X \in Pa(Y)\) so it is true. Assume it is true for all \(X, Y, Z\) such that \#\(Z > n\) and that \(\neg(X \rightarrow Y\mid Z)\) with \#\(Z = n\). Then \(Pa(Y) \notin Z\) and, by \(A_4\) and \(A_7, \neg(X \rightarrow P\mid ZYS)\) for some \(P \in Pa(Y) \setminus Z\) and \(S = Pa(Y) \setminus P\). Thus, by
hypothesis, there is a path \( X \sim P \) not intercepted by \( ZYS \). Extend the path with the link \((P,Y)\).

**Example 1:** Let \( \mathcal{X} = \{X_1, \ldots, X_4\} \) and \( \Gamma \) be a maximal consistent set containing:

\[
(X_3 \rightarrow X_4 \rightarrow X_1), \quad \neg(X_2 \rightarrow X_4),
\]
\[
(X_1 \rightarrow X_3 \rightarrow X_4), \quad \neg(X_4 \rightarrow X_1),
\]
\[
(X_4 \rightarrow X_1), \quad \neg(X_2 \rightarrow X_1), \quad \neg(X_1 \rightarrow X_4),
\]
\[
(X_1 \rightarrow X_2 \rightarrow X_4), \quad \neg(X_3 \rightarrow X_4), \quad \neg(X_1 \rightarrow X_4).
\]

Then, \( G(\Gamma) \) is the graph:

![Graph](image)

Suppose \((X_2 \rightarrow X_4 \rightarrow X_1)\), then by \((X_1 \rightarrow X_2 \rightarrow X_4)\) and \(A5, (X_1 \rightarrow X_4)\) which is a contradiction with \(\neg(X_1 \rightarrow X_4)\). Thus, \((X_2 \rightarrow X_4 \rightarrow X_1) \in \Gamma\). By Theorem 3, there must be a path \(X_2 \sim X_4\) not intercepted by \(X_1\). Fact that is clearly true in above graph.

We do not know if \(AX_{uniq} \) is a complete system for \(\mathcal{T}_{uniq}\). The next two sections present sound and complete axiomatizations for \(\mathcal{T}_{sec}\) and \(\mathcal{S}_{sec}\) respectively.

### 3.1 Axiomatization for \(\mathcal{T}_{sec}\)

Let \( \Gamma \) be a consistent set with respect to \(AX_{uniq}\). We say that an \(AX_{uniq}\)-extension \(\Gamma' \supseteq \Gamma \) is strong-recursive if and only if the graph \(G(\Gamma')\) is a DAG. Consider the following rule of inference.

**Definition 3 (Strong Recursive Inference)**

A formula \( \varphi \) can be derived from the set \( \Gamma \) by the Strong Recursive Inference (SRI) if and only if \( \varphi \) holds in all strong-recursive extensions of \( \Gamma \).

Note that since the number of possible extensions is finite, checking whether a Strong Recursive Inference is valid or not is a syntactic and decidable criterion.

For example if \(\neg(X \rightarrow Y[0]) \in \Gamma\), then (by Theorem 3) there is a path \(X \sim Y\) in any extension of \(\Gamma\). Thus, by SRI and the same Theorem, \((Y \rightarrow X[0])\) holds in any strong-recursive extension of \(\Gamma\). Thus, the SRI rule captures some notion of paths in graphs. However, it seems to be a little complex. So, do we need it? Strictly speaking, the answer is no since once \(X'\) is fixed, then all possible SRI inferences can be encoded into a single (but very large) axiom. In the other hand, it makes things easier since we don’t have to find which axiom replaces the SRI rule. We define the axiomatic system \(AX_{sec}\) as the Axioms A1–A9 plus MP and SRI as the rules of inference. Our goal is to prove that it is a sound and complete system with respect to \(\mathcal{T}_{sec}\).

**Theorem 4** \(AX_{sec}\) is sound with respect to \(\mathcal{T}_{sec}\).

**Proof:** Because of Theorem 2, we only need to prove that the SRI rule is sound with respect to \(\mathcal{T}_{sec}\). But this is true since if \(T \in \mathcal{T}_{sec}\) then Th \(T\) is a strong-recursive extension for any \(\Gamma\) such that \(T \vdash \Gamma\).

To prove completeness, it is enough to show that for every consistent \(\Gamma\) there exists a model \(T \in \mathcal{T}_{sec}\) such that \(T \vdash \Gamma\). To see this, assume the claim is true and suppose that \(\varphi\) is not provable from \(\Gamma\). Then, \(\Gamma \cup \{\neg \varphi\}\) is consistent so, by the claim, there is a model \(T\) that satisfies \(\Gamma\) and \(\neg \varphi\); i.e., \(\Gamma \not\models \varphi\). Therefore, \(\Gamma \not\models \varphi \Rightarrow \Gamma \not\models \varphi\) (equiv. to completeness) is true.

### Completeness

Let \( \Gamma \) be a maximal \(AX_{sec}\)-consistent set. We build a model for it by putting together small models whose graphs are called fragments. Each one of these models is built in order to satisfy a different negative literal \(\neg(X \rightarrow Y[Z])\) in \(\Gamma\). Formally,

**Definition 4 (Fragment)** Let \(X, Y\) be two variables and \(Z \subseteq X\). A single-connected DAG\(^3\) \(G\) with vertex set \(X\) is called an \((X, Y[Z])\)-fragment when for all \(U, V \in G\) and \(S, T \subseteq X\)

\[(i) \text{ if } \neg(X \rightarrow Y[Z]), \text{ then } G \text{ has a path } X \sim Y \text{ not intercepted by } Z,
\]
\[(ii) \text{ if } U \in G, \text{ then } U \text{ is in } X \sim Y \text{ or } U \text{ is a root},
\]
\[(iii) \text{ if } U \rightarrow V[T], \text{ then } T \text{ intersects } U \sim V, \text{ and}
\]
\[(iv) \text{ if } S \text{ is minimal such that } U \rightarrow V[S[T]), \text{ then either } T \text{ intersects } U \sim V, \text{ or } S \text{ contains an ancestor of } V \text{ not in } U \sim V.
\]

**Example 2:** Intuitively, an \((X, Y[Z])\)-fragment is a subgraph from \(G(\Gamma)\) that witnesses the literal \(\neg(X \rightarrow Y[Z])\). Let’s consider the literal \(\neg(X_1 \rightarrow X_4[0])\) in Example 1. We show that the following graph is an \((X_1, X_4[0])\)-fragment:

![Graph](image)

Clearly, it satisfies \((i)\) and \((ii)\) of Definition 4. For \((iii)\), note that the only relevances that need to be considered are \((X_1 \rightarrow X_4[3])\) and \((X_2 \rightarrow X_4[3])\) but \(X_3\) intercepts the paths \(X_2 \sim X_4\) and \(X_2 \sim X_4\). For \((iv)\), need to consider \((X_1 \rightarrow X_2 X_4[0])\). Since \(X_2\) is an ancestor of \(X_4\) not in \(X_1 \sim X_4\) then \((iv)\) is true. Therefore, the graph is a \((X_1, X_4[0])\)-fragment.

\(^3\)That is, the undirected graph associated with \(G\) has no cycles.
Theorem 5 Let $\Gamma$ be a maximal consistent set with respect to $AX_{\text{rec}}$. Then, $G(\Gamma)$ contains a $(X,Y|ZW)$-fragment for all variables $X,Y \in \mathcal{X}$ and set $Z \subseteq \mathcal{X}$.

Lemma 6 If $\neg(X \rightarrow Y|ZW)$ and $\neg(X \rightarrow YW|Z)$, then there is an $(X,Y|ZW)$-fragment where $W$ has no $Y$ ancestors.

The proofs of Theorem 5 and Lemma 6 are in the full paper [1]. The next theorem is the main result of this section.

Theorem 7 $AX_{\text{rec}}$ is complete with respect to $T_{\text{rec}}$.

Proof: Let $\Gamma$ be maximal $AX_{\text{rec}}$-consistent. Let $U,V \in \mathcal{X}$ and maximal $T \subseteq \mathcal{X}$ such that $\neg(U \rightarrow V|T)$. Let $G \leq G(\Gamma)$ be the $(U,V|T)$-fragment given by Theorem 5. Define the recursive model $T_G$ with singleton $U_G$ and $D_X$ equal to the non-negative integers for all $X \in \mathcal{X}$, by

$$Y = \begin{cases} 
0 & \text{if there is } P \text{ s.t. } P \in Pa_G(Y), P = 0, \\
\max\{X : X \in Pa_G(Y)\} & \text{otherwise.}
\end{cases}$$

This equation depends on the fixed $G$ so we write $Y^G$ when this dependency needs to be shown. By construction and properties (i)-(iii) of fragments, $T_G \vdash \neg(U \rightarrow V|T)$ and $T_G \vdash \neg(U \rightarrow V|T)$ if $(U \rightarrow V|T)$.

Let $U$ be the collection of all fragments, and $T$ the causal model with domain $U$ and equations $Y(;G) = Y^G(;)$ for $G \in U$. In the rest, we show using induction that $T$ satisfies all literals in $\Gamma$; i.e., $T \vdash \Gamma$.

First consider all literals whose first and second components are single variables. Let $X,Y \in \mathcal{X}$ and $Z \subseteq \mathcal{X}$. If $\neg(X \rightarrow Y|Z)$, then there is an $(X,Y|Z)$-fragment $G$ such that $X \rightarrow Y$ is not intercepted by $Z$. By construction, $T_G \vdash \neg(X \rightarrow Y|Z)$ which implies $T \vdash \neg(X \rightarrow Y|Z)$. Assume now $(X \rightarrow Y|Z)$ and let $G$ be any $(U,V|T)$-fragment. By definition, $Z$ intercepts all paths $X \rightarrow Y$ in $G$, so by construction $T_G \vdash \neg(X \rightarrow Y|Z)$. Thus, $T \vdash \neg(X \rightarrow Y|Z)$.

The next step is to show that $T$ satisfies all literals whose first component is a single variable. Let $(X \rightarrow YW|Z)$ be a such negative literal with $Y \in \mathcal{X}, W \subseteq \mathcal{X}$. Without loss of generality assume that $\neg(X \rightarrow Y|ZW)$. By Lemma 6, there is a fragment $G$ such that $ZW$ doesn’t intercepts $X \rightarrow Y$ and $W$ contains no ancestor of $Y$. By construction, $T_G \vdash \neg(X \rightarrow YW|Z)$. Assume now $(X \rightarrow YW|Z)$ and, without loss of generality that $\neg(X \rightarrow Y|ZW)$ since otherwise $T \vdash (X \rightarrow YW|Z)$ (the reader can check it using A4). Let $G$ be an $(U,V|T)$-fragment such that $Z$ doesn’t intercept path $X \rightarrow Y$. Then, (iv)

| of Definition 4 and construction imply $T_G \vdash (X \rightarrow YW|Z)$. Thus, $T \vdash (X \rightarrow YW|Z)$.

The last step is for general literals. Note that A3 and A6 imply $(XW \rightarrow Y|Z) \Leftrightarrow (X \rightarrow Y|ZW) \& (W \rightarrow Y|ZX)$. Therefore, by above cases, $(XW \rightarrow Y|Z)$ if and only if $T \vdash (XW \rightarrow Y|Z)$.

Example 3: Consider the $(X_1,X_4|\emptyset)$-fragment $G$ in Example 2. The construction of Theorem 7 gives the model $T_G$:

$$
\begin{align*}
X_1 &= 0, \\
X_2 &= 0, \\
X_4 &= X_3, \\
X_3 &= \begin{cases} 
0 & \text{if } X_1 = 0 \text{ or } X_2 = 0, \\
\max\{X_1,X_2\} & \text{otherwise.}
\end{cases}
\end{align*}
$$

We now prove that $T_G$ satisfies $\neg(X_1 \rightarrow X_4|\emptyset)$ and all positive literals in Example 1. It is easy to check that $[X_1 = 0, X_2 = 1]X_4 = 0$ and $[X_1 = 1, X_2 = 1]X_4 = 1$. Thus, $\neg(X_1 \rightarrow X_4|X_2)$ which implies $\neg(X_1 \rightarrow X_4|\emptyset)$. Clearly, the model satisfies $(X_4 \rightarrow X_3|\emptyset)$, $(X_1X_2X_4 \rightarrow X_2|\emptyset)$ and $(X_2X_3X_4 \rightarrow X_1|\emptyset)$. It also satisfies $(X_1X_2 \rightarrow X_3|X_3)$ since the equation for $X_4$ depends only in $X_3$. We need to show that it satisfies $(X_1 \rightarrow X_2X_4|\emptyset)$:

$$[x_1]X_4 = [x_1,[x_1]X_2]X_4$$

which is independent of $x_1$. Thus, $T_G \vdash (X_1 \rightarrow X_2X_4|\emptyset)$.

3.2 Axiomatization for $T_{\text{rec}}$

Let $\Gamma$ be a consistent set with respect to $AX_{\text{uniq}}$. We say that an $AX_{\text{uniq}}$-extension $\Gamma' \supseteq \Gamma$ is recursive if and only if the graph $G'(\Gamma')$ contains an $(X,Y|Z)$-fragment for all $X,Y \in \mathcal{X}$ and $Z \subseteq \mathcal{X}$. Similar to SRI, we make

Definition 5 (Recursive Inference) A formula $\varphi$ can be derived from set $\Gamma$ by Recursive Inference (RI) if and only if $\varphi$ holds in all recursive extensions of $\Gamma$.

As with SRI, all RI inferences are decidable. Let $AX_{\text{rec}}$ be the axiomatic systems defined as the Axioms A1–A9 plus MP and RI as the rules of inference. Our goal is to prove that it is sound and complete for $T_{\text{rec}}$.

Theorem 8 $AX_{\text{rec}}$ is sound with respect to $T_{\text{rec}}$.

Proof: Because of Theorem 2, we only need to prove that the RI rule is sound. Let $T \in T_{\text{rec}}$ and $\Gamma$ such that $T \vdash \Gamma$. Let $X,Y \in \mathcal{X}$ and $Z \subseteq \mathcal{X}$. If

[4] An empty max is defined to be 0.
(X → Y | Z), then the empty graph is a (X, Y | Z)-fragment. If \( \neg(X \rightarrow Y | Z) \), then there is \( u \in \mathcal{U} \) such that \( T(u) \models \neg(X \rightarrow Y | Z) \). By Theorem 5, \( G(Th T(u)) \) contains a (X, Y | Z)-fragment. Then, \( Th T \) is a recursive extension of \( \Gamma \).

Let \( T_1 \) and \( T_2 \) be two functional causal models with exogenous domains \( U_1 \), \( U_2 \), and equations \( F_Y^i \) for \( i = 1, 2 \) and \( Y \in X \). The direct sum of \( T_1 \) and \( T_2 \), denoted by \( T_1 \oplus T_2 \), is defined as the causal model \( T \) with exogenous domain \( \mathcal{U} = \{1\} \times U_1 \cup \{2\} \times U_2 \) and equations given by \( F_Y(\cdot; i, u) = F_Y^i(\cdot; u) \) for \( i = 1, 2 \). Abusing notation, we say that the causal model \( T_1 \oplus \cdots \oplus T_n \) is the model with exogenous domain \( \mathcal{U} = \bigcup_i \{i\} \times U_i \) and equations given by \( F_Y(\cdot; i, u) = F_Y^i(\cdot; u) \) for \( 1 \leq i \leq n \).

Theorem 9 Let \( T, T_1, \ldots, T_n \in \mathcal{T}_{uniq} \) be causal models. Then,

\[ \bigcup_{i=1}^n \text{Lit}^-(Th T_i) \cup \bigcap_{i=1}^n \text{Lit}^+(Th T_i). \tag{3} \]

(i) \( T \cong T_1 \oplus \cdots \oplus T_n \) if and only if \( Th T \) is the unique AX_{uniq}-extension of

\[ \bigcup_{i=1}^n \text{Lit}^-(Th T_i) \cup \bigcap_{i=1}^n \text{Lit}^+(Th T_i). \]

(ii) If \( T \cong T_1 \oplus \cdots \oplus T_n \), then \( G(T) = G(T_1) \cup \cdots \cup G(T_n) \).

Proof: (i) First we prove that \( \Gamma \) has only one extension. Let \( \varphi \in \mathcal{C} \) be a positive literal. If \( \exists i[T_i \models \varphi], \) then \( \exists i[T_i \models \neg \varphi] \) which implies \( \neg \varphi \in \Gamma \). If \( \forall i[T_i \models \varphi], \) then \( \varphi \in \text{Lit}^+(Th T_1) \) which implies \( \varphi \in \Gamma \). For negative literals consider \( \psi = \neg \varphi \). Thus, \( \Gamma \) is a maximal consistent set, so its unique extension is \( Th_{uniq} \).

Assume \( T \cong T_1 \oplus \cdots \oplus T_n \), and \( \varphi \) a positive literal. If \( T \models \varphi \), then \( \Gamma_{uniq}[T(i, u) \models \varphi] \) which implies \( \forall u[T(i, u) \models \varphi] \). If \( T \not\models \varphi \), then \( \exists (i, u)[T(i, u) \not\models \varphi] \) which implies \( \exists \varphi \not\in \text{Lit}^-(Th T_i)] \), so \( \varphi \in \Gamma \). Therefore, \( \text{Lit}(Th T) = \text{Lit} \Gamma \) which implies the result.

Now, assume \( Th T = Th_{uniq} \). Let \( \varphi \) be a positive literal. Then, \( \varphi \in Th T \) iff

\[ \varphi \in \bigcup_{i=1}^n \text{Lit}^+(Th T_i) \iff \varphi \in Th (T_1 \oplus \cdots \oplus T_n). \]

(iii) Assume \( T \cong T_1 \oplus \cdots \oplus T_n \). Let \( X, Y \) be two variables and \( W = \{X, Y\} \). Then, \( (X, Y) \in G(Th (T_1 \oplus \cdots \oplus T_n)) \) iff \( \exists i[T_i \models \neg(X \rightarrow Y | W)] \) iff \( \exists i[(X, Y) \in G(T_i)] \).

Completeness

To motivate our proof for completeness, observe that a model \( T \in \mathcal{T}_{rec} \) is composed of different \( \mathcal{T}_{rec} \) models (one per \( u \in \mathcal{U} \)), and that each negative literal satisfied by \( T \) is satisfied by one of those models. We will do something similar but with sets of formulas. Thus, to each maximal consistent set \( \Gamma \) and negative literal \( \varphi \in \Gamma \), we assign a maximal AX_{rec}-consistent set \( \Gamma' \subseteq \Gamma \) containing \( \varphi \). All such sets are then used to build \( \mathcal{T}_{rec} \) models that combine into a model \( T \in \mathcal{T}_{rec} \) such that \( T \models \Gamma \).

Let \( \psi_0, \psi_1, \ldots, \psi_n \) be a fixed enumeration of all literals in \( L \). Fix a maximal AX_{rec}-consistent set \( \Gamma \). Let \( X, Y \) be variables, \( \varphi = \neg(X \rightarrow Y | Z) \) a negative literal in \( \Gamma \), and define the sequence of sets \( \{\psi_i\}_{i=1}^n \) as

\[ \Psi_0 = Th_{rec}(\{\varphi\} \cup \text{Lit}^+ \Gamma) \]

\[ \Psi_{i+1} = \begin{cases} Th_{rec}(\psi_i \cup \{\psi_i\}) & \text{if } \psi_i \cup \{\psi_i\} \text{ is AX}_{rec}-consistent, \\ Th_{rec}(\psi_i \cup \{-\psi_i\}) & \text{otherwise}. \end{cases} \]

We want to prove that each \( \psi_i \) is AX_{rec}-consistent. By construction, it is enough to show that \( \Psi_0 \) is consistent. Fact that is true since \( \Gamma \) is AX_{rec}-consistent; i.e., \( G(\Gamma) \) contains an \( (X, Y | Z) \)-fragment for \( \varphi \). Note that \( \Psi_n \) is maximal and that \( \text{Lit}^- \Psi \subseteq \text{Lit}^- \Gamma \). We call \( \Psi_n \) a \( (\varphi, \Gamma) \)-foliation.

Theorem 10 AX_{rec} is complete with respect to \( \mathcal{T}_{rec} \).

Proof: Let \( \Gamma \) be a maximal AX_{rec}-consistent, and \( \Psi \) a \( (\varphi, \Gamma) \)-foliation. Denote with \( Th\Psi \) the strong recursive model such that \( T\Psi \models \Psi \) (it exists by completeness of AX_{rec}). By a simple induction in the structure of negative literals (similar to that in proof of Theorem 7), it can be shown that for all literal \( \neg(U \rightarrow Z | T) \) in \( \Gamma \), there exists a foliation \( \Psi \) such that \( Th\Psi \models \neg(U \rightarrow Z | T) \). Let \( \mathcal{F} = \{ \Psi \vdash \psi \models (\varphi, \Gamma) \}-foliation \), and \( T = \bigcup_{\psi \in \mathcal{F}} T\Psi \). Then, by Theorem 9, \( T \models \Gamma \).

Now, we prove the claim made in diagram (2) about the subclasses of causal models.

Theorem 11 \( \mathcal{T}_{rec} < \mathcal{T}_{uniq} \).

Proof: Consider the model \( T \) with variables \( \{X_1, \ldots, X_4\} \), singleton \( U \) and equations
\[ X_1 = 1, \quad X_2 = (X_1, X_3), \quad X_3 = X_2 X_4, \quad X_4 = X_2 X_4. \]

It is a non-recursive model that satisfies \( \neg(X_1 \rightarrow X_4 | \emptyset) \) and \( (X_1 \rightarrow X_4 | Z) \) for \( Z \in \{X_2, X_3\} \). Note that \( G(T) = G(Th T) \) is the graph with edges \( \{X_1 \rightarrow X_2 \rightarrow X_4, X_2 \rightarrow X_3\} \), and that \( T \) satisfies \( \neg(X_2 \rightarrow X_4 | X_1 X_3), \neg(X_1 \rightarrow X_2 | X_2 X_3), \neg(X_2 \rightarrow X_2 | X_3 X_4) \). It can be shown that \( Th T \) has no recursive extension, so it is not AX_{rec}-consistent. Therefore, there is no recursive model \( L \)-equivalent to it.
4 Applications

In this section we consider the question of how to use information about causal relevance for the problem of identifying causal structure (also known as the problem of discovering causal structure). Fix a model $T$ and let $\Gamma$ be a set of causal relevances that hold in $T$; i.e., $T \models \Gamma$. If $T \in \mathcal{T}_{\text{rec}}$, then the completeness of $\mathcal{A}_{\text{rec}}$ and Theorem 3 imply $G(T)$ is equal to $G(\Gamma')$ for some $\mathcal{A}_{\text{rec}}$-extension $\Gamma'$ of $\Gamma$. Thus, the graph defined as

$$G(\Gamma) \overset{\text{def}}{=} \bigcap_{\Gamma'} G(\Gamma'),$$

where the intersection is over all $\mathcal{A}_{\text{rec}}$-extensions of $\Gamma$, is an edge subgraph of $G(T)$. Completeness is needed here in order to guarantee that every $\Gamma'$ has a corresponding model in $\mathcal{T}_{\text{rec}}$. Suppose further that we can buy one of several possible information options $\Gamma_i \supseteq \Gamma$. The problem is to decide which is the best option. Let $c_i$ be the cost associated to $\Gamma_i$. A rational approach would be to buy the set that minimizes the pondered cost $c_i \overset{\text{def}}{=} c_i / \#(G(\Gamma_i) - G(\Gamma))$ where $\#G$ denotes the number of edges in graph $G$. The formula for $c_i$ considers that the benefit of the information is proportional to the number of new edges provided by it. Note that if $\Gamma_i$ generates no new edges, then $c_i = \infty$.

Similar definitions can be made when $T \in \mathcal{T}_{\text{uniq}}$. Observe that we do not claim this can be implemented in an efficient way. Nonetheless, this calculus can be integrated into other methods of structure identification by using it to prune the search space; e.g., once we know $\neg(X \rightarrow Y | Z)$, then the graph we are looking for must have a path $X \rightarrow Y$ not intercepted by $Z$.

Another application of the calculus is for testing whether a model is recursive or not. Since $\mathcal{A}_{\text{rec}}$ is complete with respect to $\mathcal{T}_{\text{rec}}$, then a necessary test for recursiveness is $\mathcal{A}_{\text{rec}}$-consistency. Thus, if $\Gamma$ is not $\mathcal{A}_{\text{rec}}$-consistent, then $\Gamma$ corresponds to a non-recursive model. An argument like this one was used in the proof of Theorem 11.

5 Summary

There are three main contributions in this paper. First, it gives a finer decomposition $\mathcal{T}_{\text{uniq}}$ into the two subclasses $\mathcal{T}_{\text{rec}}$ and $\mathcal{T}_{\text{uniq}}$. This decomposition is necessary for the formalization of causal relevance from the three-place relation $(X \rightarrow Y | Z)$, and it is useful as shown in the applications of the calculus. The second contribution is the collection of axiomatizations $\mathcal{A}_{\text{uniq}}, \mathcal{A}_{\text{rec}}$ and $\mathcal{A}_{\text{rec}}$. $\mathcal{A}_{\text{uniq}}$ extends the system in [6] with the three new axioms A5, A8 and A9. It is a sound axiomatization with respect to the class $\mathcal{T}_{\text{uniq}}$. $\mathcal{A}_{\text{rec}}$ is a stronger version of $\mathcal{A}_{\text{uniq}}$ that is sound and complete with respect to $\mathcal{T}_{\text{rec}}$. It is equal to $\mathcal{A}_{\text{uniq}}$ plus the rule RI. The axiomatization $\mathcal{A}_{\text{rec}}$, a stronger version $\mathcal{A}_{\text{rec}}$, is sound and complete with respect to $\mathcal{T}_{\text{rec}}$. It is equal to $\mathcal{A}_{\text{uniq}}$ plus the rule SRI. Finally, we show how the calculus of causal relevance can be used to identify causal structure. This is a non-probabilistic approach that can be used to infer structure from non-observational data. All other methods we know of are based on probabilistic inference (e.g., [9]). It would be interesting to study the similarities and difference between these approaches.

Some open questions are if $\mathcal{A}_{\text{uniq}}$ is a complete system with respect to $\mathcal{T}_{\text{uniq}}$, if the collection $\mathcal{A}_1-\mathcal{A}_9$ is an independent set of Axioms, and which Axioms can be used to replace the RI and SRI rules.

Acknowledgments

Thanks to Carlos Brito and Judea Pearl for their useful comments in a first version of this paper. They spotted a number of errors, omissions and ways to improve it. Thanks also Jin Tian and people at the AI seminar at UCLA.

References

[1] B. Bonet (2001). A Calculus for Causal Relevance. Cognitive System Lab. Technical Report R-268. UCLA.
[2] D. Galles and J. Pearl (1997). Axioms of Causal Relevance. Artificial Intelligence 97(1-2), pp. 9-43.
[3] D. Geiger, T. Verma and J. Pearl (1990). Identifying Independence in Bayesian Networks. Networks 20, pp. 507-534.
[4] J. Halpern (1998). Axiomatizing Causal Reasoning. In Proceedings UAI-98, pp. 202-210. Morgan Kaufmann.
[5] J. Pearl and A. Paz (1987). Graphoids: A graph-based logic for reasoning about relevance relations. Advances in Artificial Intelligence, vol. II. North-Holland.
[6] J. Pearl (2000). Causality: Models, Reasoning, and Inference. Cambridge University Press.
[7] W.C. Salmon (1984). Scientific Explanation and the Causal Structure of the World. Princeton University Press.
[8] P. Suppes (1970). A Probabilistic Theory of Causality. North-Holland.
[9] J. Tian and J. Pearl (2001). Causal Discovery from Changes. In Proceedings UAI-2001. Morgan Kaufmann.