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 func-
tional causal models as semantics. The cal-
culus consists of axioms and rules of inference
for reasoning about causal relevance relation-
ships. 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 spe-
cific 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 chang-
ing X alter Y?" Formal interest in this notion ap-
pears 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 in-
ference. 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 Sec-
tion 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. Sec-
tion 3 presents the axiomatizations, some general re-
sults about the calculus as well as the proofs for sound-
ness and completeness. Then, Section 4 shows how
this theory can be used to solve identification prob-
lems. We conclude with a brief summary.

2 Model Theory

A causal model is a set of variables in which each vari-
able may have influence on others. The variables are
divided in two groups. The *endogenous* variables for
which the model provides a description of the mech-
anisms that influence them, and the *exogenous* vari-
ables 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 *modifiable structural equations*. Each equation is a function

$$F_Y : \prod_{\substack{X \in \mathcal{X} \\ X \neq Y}} \mathcal{D}_X \times \mathcal{U} \longrightarrow \mathcal{D}_Y,$$

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 $[x]T$, that results from the *intervention* of setting the value of X to x, as $[x]T = (\mathcal{U}, \{F'_Y : Y \in \mathcal{X}\})$ where $F'_Y = F_Y$ for $Y \notin X$ and $F'_Y = x{\downarrow}_Y$ for $Y \in X$. This intervention is denoted by the operator $do(X = x)$ that maps causal models into causal models by $T \rightsquigarrow [x]T$ (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 $[x]T$ 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 *potential response* for variable Y and $u \in \mathcal{U}$ under intervention $do(X = x)$ is defined as the solution Y in model $[x]T$ when the exogenous variable has value u (denoted by the function $u \rightsquigarrow [x]Y(u)$). The main difference between causal models and Bayesian networks is that the former incorporate a semantics for all interventions $do(X = x)$.

A *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.

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

Composition can be briefly summarized by the equation $[x]Y(u) = [x, [x]W(u)]Y(u)$ which is valid for all variables Y, W, all valuations x of X, 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 *previously recorded* value from another intervention.

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 $\mathcal{X}$, is denoted by $\mathcal{L}$. Each atomic formula, or just

atom, of $\mathcal{L}$ is of the form $(X \nrightarrow Y | Z)$ where X, Y and Z are *disjoint* subsets of $\mathcal{X}$, and X, Y are non-empty. Intuitively, an atom $(X \nrightarrow Y | Z)$ is read as *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 *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 Γ is a set of formulas, the set of positive (resp. negative) literals in Γ is denoted by $\text{Lit}^+ \Gamma$ (resp. $\text{Lit}^- \Gamma$).

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.

Definition 1 *A causal model $T \in \mathcal{T}_{\text{uniq}}$ satisfies the atom $(X \nrightarrow Y | Z)$ if and only if*

$$[x, z, w]Y(\cdot) = [x', z, w]Y(\cdot) \tag{1}$$

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 φ, and $T \vDash \Gamma$ when T satisfies all formulas in Γ. The interpretation is extended to all formulas by *(i)* $T \vDash \neg\varphi$ iff $T \nvDash \varphi$, *(ii)* $T \vDash (\varphi \,\&\, \psi)$ iff $T \vDash \varphi$ and $T \vDash \psi$, *(iii)* $T \vDash (\varphi \vee \psi)$ iff $T \vDash \varphi$ or $T \vDash \psi$, and *(iv)* $T \nvDash (\varphi \Rightarrow \psi)$ iff $T \vDash \varphi$ but $T \nvDash \psi$.

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

2.4 Classes of Models

We begin with some definitions. The *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 $G \leq G'$ when G is an edge subgraph of G'.

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

Let $\mathcal{T}_{\text{srec}}$ and $\mathcal{T}_{\text{rec}}$ be the subclasses of models from $\mathcal{T}_{\text{uniq}}$ that are strong recursive and recursive respectively. By definition, $\mathcal{T}_{\text{srec}} \subseteq \mathcal{T}_{\text{rec}} \subseteq \mathcal{T}_{\text{uniq}}$.

We say that two models T, T' are $\mathcal{L}$–equivalent, written $T \cong T'$, if they can not be distinguished by the language; i.e., $\forall\varphi[T \vDash \varphi \Leftrightarrow T' \vDash \varphi]$. We say that a class $\mathcal{T}$ is $\mathcal{L}$-contained into another class $\mathcal{T}'$, written $\mathcal{T} \leq \mathcal{T}'$, if $\forall(T \in \mathcal{T})\exists(T' \in \mathcal{T}')[T \cong T']$. The following diagram shows the partial order induced by $\leq$ among the classes of models (modulo $\cong$). The labels at the arrows indicate whether the relation is strict or not.

$$\mathcal{T}_{\text{srec}} \xrightarrow{\ <\ } \mathcal{T}_{\text{rec}} \xrightarrow{\ <\ } \mathcal{T}_{\text{uniq}} \qquad (2)$$

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

$$X_1(u) = X_2(u), \qquad X_1(u') = 0,$$
$$X_2(u) = 0, \qquad X_2(u') = X_1(u') \ ;$$

i.e., with graph $X_1 \rightleftarrows X_2$. 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 $\cong$. To simplify notation, however, we write T and $\mathcal{T}$ instead of $[T]_{\cong}$ and $\mathcal{T}/\cong$ 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 *license* that allows us to derive new formulas from previous ones. A *proof* for formula φ from a set Γ is a finite sequence of formulas $\varphi_1, \varphi_2, \dots, \varphi_n = \varphi$ such that each φ_i is either an axiom, a formula in Γ, or a formula derived using a rule of inference from previous ones. We write $\Gamma \vdash_{\text{AX}} \varphi$ if φ can be proved from Γ in the AX system. An axiomatic system AX is *sound* with respect to a class of models $\mathcal{T}$ iff $\Gamma \vdash_{\text{AX}} \varphi$ implies $\Gamma \vDash_{\mathcal{T}} \varphi$;[1] i.e., everything that can be proved from Γ is valid in all models that satisfy Γ. AX is *complete* with respect to $\mathcal{T}$ iff $\Gamma \vDash_{\mathcal{T}} \varphi$ implies $\Gamma \vdash_{\text{AX}} \varphi$; i.e., everything that is valid in models that satisfy Γ can be proved from Γ.

The following list contains all the axioms considered in this paper. It is assumed that the variables in the formulas make them valid $\mathcal{L}$-formulas; e.g., that W is disjoint of XYZ in A2.[2]

A1. All propositional tautologies,

A2. *Strong Union:*
$(X \nrightarrow Y|Z) \Rightarrow (X \nrightarrow Y|ZW),$

A3. *Left Decomposition:*
$(XW \nrightarrow Y|Z) \Rightarrow (X \nrightarrow Y|Z),$

A4. *Weak Right Decomposition:*
$(X \nrightarrow YW|Z) \ \& \ (X \nrightarrow Y|ZW) \Rightarrow (X \nrightarrow Y|Z),$

A5. *Weak Transitivity:*
$(X \nrightarrow YW|Z) \ \& \ (W \nrightarrow Y|ZX) \Rightarrow (X \nrightarrow Y|ZW),$

A6. *Left Intersection:*
$(X \nrightarrow Y|ZW) \ \& \ (W \nrightarrow Y|ZX) \Rightarrow (XW \nrightarrow Y|Z),$

A7. *Right Intersection:*
$(X \nrightarrow Y|ZW) \ \& \ (X \nrightarrow W|ZY) \Rightarrow (X \nrightarrow YW|Z),$

A8. *Linearity:*
$(X \nrightarrow Y|ZV) \ \& \ (X \nrightarrow Y|ZU) \ \& \ (U \nrightarrow V|ZXW) \ \&$
$(V \nrightarrow U|ZXW) \Rightarrow (X \nrightarrow Y|ZW),$

A9. *Context Substitution:*
$(X \nrightarrow YW|Z) \ \& \ (X \nrightarrow V|ZYW) \ \& \ (W \nrightarrow Y|ZX)$
$\Rightarrow (X \nrightarrow YV|Z).$

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

Let $\Gamma \subseteq \mathcal{L}$ be a set of formulas and fix an axiomatic system AX. The following definitions are standard. Γ is a *closed* set if $\forall\varphi[\Gamma \vdash_{\text{AX}} \varphi \Rightarrow \varphi \in \Gamma]$. The *theory* associated with Γ, denoted as $\text{Th}_{\text{AX}} \Gamma$, is the smallest closed set containing Γ. Γ is a *consistent* set if $\forall\varphi[\varphi \in \text{Th}_{\text{AX}} \Gamma \Rightarrow \neg\varphi \notin \text{Th}_{\text{AX}} \Gamma]$. A set $\Gamma' \supseteq \Gamma$ is an *extension* of Γ if it is consistent and maximal. The theory associated with a model T is the set of all formulas valid in T; i.e. $\text{Th}\, T = \{\varphi \in \mathcal{L} : T \vDash \varphi\}$.

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

Theorem 2 AX_{uniq} *is sound with respect to* $\mathcal{T}_{\text{uniq}}$.

Proof: The soundness of A1 is obvious. Fix $u \in \mathcal{U}$. For clarity, we won't write the dependency in u. 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' will be valuations for X. In each case, the suppositions are denoted from first to last by S1, S2, ...

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

$$
\begin{aligned}
[x, z, v]Y &= [x, z, v, [x, z, v]W]Y && \text{(comp.)}\\
&= [x', z, v, [x, z, v]W]Y && \text{(by S2)}\\
&= [x', z, v, [x', z, v]W]Y && \text{(by S1)}\\
&= [x', z, v]Y. && \text{(comp.)}
\end{aligned}
$$

Thus, $(X \nrightarrow Y|Z)$.

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

$$
\begin{aligned}
[x,z,v,w]Y &= [x,z,v,[x,z,v]W]Y &\text{(by S2)}\\
&= [x,z,v]Y &\text{(comp.)}\\
&= [x',z,v]Y &\text{(by S1)}\\
&= [x',z,v,[x',z,v]W]Y &\text{(comp.)}\\
&= [x',z,v,w]Y. &\text{(by S2)}
\end{aligned}
$$

Thus, $(X \nrightarrow Y|ZW)$.

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

$$
\begin{aligned}
[x,w,z,t]Y&\\
&= [x,w,z,t,[x,w,z,t]U]Y &\text{(comp.)}\\
&= [x',w,z,t,[x,w,z,t]U]Y &\text{(by S2)}\\
&= [x',w,z,t,[x,w,z,t]U, &\text{(comp.)}\\
&\quad\ [x',w,z,t,[x,w,z,t]U]V]Y\\
&= [x',w,z,t,[x,w,z,t]U, &\text{(by S3)}\\
&\quad\ [x',w,z,t,[x',w,z,t]U]V]Y\\
&= [x',w,z,t,[x,w,z,t]U,[x',w,z,t]V]Y &\text{(comp.)}\\
&= [x',w,z,t,[x',w,z,t]V, &\text{(comp.)}\\
&\quad\ [x,w,z,t,[x,w,z,t]V]U]Y\\
&= [x',w,z,t,[x',w,z,t]V, &\text{(by S4)}\\
&\quad\ [x,w,z,t,[x',w,z,t]V]U]Y\\
&= [x,w,z,t,[x',w,z,t]V, &\text{(by S2)}\\
&\quad\ [x,w,z,t,[x',w,z,t]V]U]Y\\
&= [x,w,z,t,[x',w,z,t]V]Y &\text{(comp.)}\\
&= [x',w,z,t,[x',w,z,t]V]Y &\text{(by S1)}\\
&= [x',w,z,t]Y. &\text{(comp.)}
\end{aligned}
$$

Thus, $(X \nrightarrow Y|WZ)$.

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

$$
\begin{aligned}
[x,z,t]V &= [x,z,t,[x,z,t]W, &\text{(comp.)}\\
&\quad\ [x,z,t,[x,z,t]W]Y]V\\
&= [x,z,t,[x,z,t]W,[x,z,t]Y]V &\text{(comp.)}\\
&= [x',z,t,[x',z,t]W,[x',z,t]Y]V &\text{(by S1)}\\
&= [x',z,t,[x',z,t]W, &\text{(comp.)}\\
&\quad\ [x',z,t,[x',z,t]W]Y]V\\
&= [x',z,t]V. &\text{(comp.)}
\end{aligned}
$$

Thus, $(X \nrightarrow YV|Z)$. ∎

Some Theorems of AX$_{\text{uniq}}$ are:

1 *Extended Left Intersection:*
$$(TX \nrightarrow Y|ZW)\&(TW \nrightarrow Y|ZX) \Rightarrow (TXW \nrightarrow Y|Z),$$

2 *Extended Right Intersection:*
$$(X \nrightarrow TY|ZW)\&(X \nrightarrow TW|ZX) \Rightarrow (X \nrightarrow TYW|Z).$$

Let Γ 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 $\text{Pa}(Y)$, if $((PY)^c \nrightarrow Y|P) \in \Gamma$ and P is $\subseteq$-minimal. If such set P does not exist, then $\text{Pa}(Y)$ is defined as Y^c. The *syntactic graph* associated with a maximal consistent set Γ is the directed graph $G(\Gamma)$ with vertex set equal to $\mathcal{X}$ and edge set E such that $(X,Y) \in E$ if and only if $X \in \text{Pa}(Y)$. From now on, if the set Γ is clear from context, we do not write the suffix "$\in \Gamma$".

Theorem 3 *Let Γ be a maximal AX$_{\text{uniq}}$-consistent set and T a causal model with exogenous domain $\mathcal{U}$. Then,*

(i) $\text{Pa}(Y)$ *is unique (i.e., $G(\Gamma)$ is well defined),*

(ii) $X \in \text{Pa}(Y)$ *iff* $\neg(X \nrightarrow Y|W)$ *for all W,*

(iii) $G(T) = G(\text{Th}\,T)$,

(iv) *if $\neg(X \nrightarrow Y|Z)$, then there is a (directed) path $X \rightsquigarrow Y$ not intercepted 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 $W_1 = (P_1Y)^c$, $W_2 = (P_2Y)^c$, $(W_1 \nrightarrow Y|P_1)$ and $(W_2 \nrightarrow Y|P_2)$. Let $P = P_1 \cap P_2$. Then, by $P_1 \setminus P \subseteq W_2$, $P_2 \setminus P \subseteq W_1$ and Extended Left Intersection, $(W \nrightarrow Y|P)$ where $W = (YP)^c$.

(ii). Let $P = \text{Pa}(Y) \setminus X$, $W = (XYP)^c$. Then, $\neg(WX \nrightarrow Y|P)$ and, by A6, $\neg(X \nrightarrow Y|WP)$. So, the only if direction follows from A2. If $\neg(X \nrightarrow Y|(XY)^c)$, then, by A2 and definition, $X \in \text{Pa}(Y)$.

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

$$
\begin{aligned}
(X,Y) &\in G(T)\\
&\iff \exists u[(X,Y) \in G(T,u)]\\
&\iff \exists u[X \text{ is non-trivial arg. of } F_Y(\,\cdot\,;u)]\\
&\iff \exists u[T(u) \vDash \neg(X \nrightarrow Y|W)]\\
&\iff \neg(X \nrightarrow Y|W) \in \text{Th}\,T\\
&\iff X \in \text{Pa}(Y)\,.
\end{aligned}
$$

(iv). By reverse induction in $\#Z$ (cardinality of Z). Let $Z = (XY)^c$ such that $\neg(X \nrightarrow Y|Z)$. Then, $X \in \text{Pa}(Y)$ so it is true. Assume it is true for all X, Y, Z such that $\#Z > n$ and that $\neg(X \nrightarrow Y|Z)$ with $\#Z = n$. Then $\text{Pa}(Y) \nsubseteq Z$ and, by A4 and A7, $\neg(X \nrightarrow P|ZYS)$ for some $P \in \text{Pa}(Y) \setminus Z$ and $S = \text{Pa}(Y) \setminus P$. Thus, by

hypothesis, there is a path $X \rightsquigarrow P$ not intercepted by ZYS. Extend the path with the link (P, Y). ∎

Example 1: Let $\mathcal{X} = \{X_1, \dots, X_4\}$ and Γ be a maximal consistent set containing:

$(X_3 X_4 \nrightarrow X_1 | \emptyset)$, $\neg(X_2 \nrightarrow X_1 | X_3 X_4)$,

$(X_1 X_3 X_4 \nrightarrow X_2 | \emptyset)$, $(X_1 \nrightarrow X_2 X_4 | \emptyset)$,

$(X_4 \nrightarrow X_3 | \emptyset)$, $\neg(X_2 \nrightarrow X_3 | X_1 X_4)$, $\neg(X_1 \nrightarrow X_3 | X_2 X_4)$,

$(X_1 X_2 \nrightarrow X_4 | X_3)$, $\neg(X_3 \nrightarrow X_4 | X_1 X_2)$, $\neg(X_1 \nrightarrow X_4 | X_2)$.

Then, $G(\Gamma)$ is the graph:
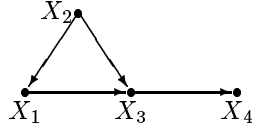
Suppose $(X_2 \nrightarrow X_4 | X_1)$, then, by $(X_1 \nrightarrow X_2 X_4 | \emptyset)$ and A5, $(X_1 \nrightarrow X_4 | X_2)$ which is a contradiction with $\neg(X_1 \nrightarrow X_4 | X_2)$. Thus, $\neg(X_2 \nrightarrow X_4 | X_1) \in \Gamma$. By Theorem 3, there must be a path $X_2 \rightsquigarrow X_4$ 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 $\mathcal{T}_{\text{uniq}}$. The next two sections present sound and complete axiomatizations for $\mathcal{T}_{\text{srec}}$ and $\mathcal{T}_{\text{rec}}$ respectively.

3.1 Axiomatization for $\mathcal{T}_{\text{srec}}$

Let Γ 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 φ can be derived from the set Γ by the Strong Recursive Inference (SRI) if and only if φ holds in all strong–recursive extensions of Γ.

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 $\neg(X \nrightarrow Y | \emptyset) \in \Gamma$, then (by Theorem 3) there is a path $X \rightsquigarrow Y$ in any extension of Γ. Thus, by SRI and the same Theorem, $(Y \nrightarrow X | \emptyset)$ holds in any strong–recursive extension of Γ. 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 $\mathcal{X}$ 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_{srec} 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 $\mathcal{T}_{\text{srec}}$.

Theorem 4 AX_{srec} *is sound with respect to* $\mathcal{T}_{\text{srec}}$.

Proof: Because of Theorem 2, we only need to prove that the SRI rule is sound with respect to $\mathcal{T}_{\text{srec}}$. But this is true since if $T \in \mathcal{T}_{\text{srec}}$, then $\text{Th}\, T$ is a strong–recursive extension for any Γ such that $T \vDash \Gamma$. ∎

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

Completeness

Let Γ be a maximal AX_{srec}–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 $\neg(X \nrightarrow Y | Z)$ in Γ. Formally,

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

(i) if $\neg(X \nrightarrow Y | Z)$, then G has a path $X \rightsquigarrow Y$ not intercepted by Z,

(ii) if $U \in G$, then U is in $X \rightsquigarrow Y$ or U is a root,

(iii) if $(U \nrightarrow V | T)$, then T intercepts $U \rightsquigarrow V$, and

(iv) if S is minimal such that $(U \nrightarrow V S | T)$, then either T intercepts $U \rightsquigarrow V$, or S contains an ancestor of V not in $U \rightsquigarrow V$.

Example 2: Intuitively, an $(X, Y | Z)$–fragment is a subgraph from $G(\Gamma)$ that witnesses the literal $\neg(X \nrightarrow Y | Z)$. Let's consider the literal $\neg(X_1 \nrightarrow X_4 | \emptyset)$ in Example 1. We show that the following graph is an $(X_1, X_4 | \emptyset)$–fragment:
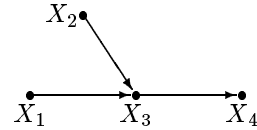
Clearly, it satisfies *(i)* and *(ii)* of Definition 4. For *(iii)*, note that the only relevances that need to be considered are $(X_1 \nrightarrow X_4 | X_3)$ and $(X_2 \nrightarrow X_4 | X_3)$ but X_3 intercepts the paths $X_1 \rightsquigarrow X_4$ and $X_2 \rightsquigarrow X_4$. For *(iv)*, need to consider $(X_1 \nrightarrow X_2 X_4 | \emptyset)$. Since X_2 is an ancestor of X_4 not in $X_1 \rightsquigarrow X_4$ then *(iv)* is true. Therefore, the graph is a $(X_1, X_4 | \emptyset)$–fragment. ∎

---

[3]That is, the undirected graph associated with G has no cycles.

Theorem 5 *Let Γ be a maximal consistent set with respect to AX_{srec}. Then, $G(\Gamma)$ contains a $(X,Y|Z)$–fragment for all variables $X,Y \in \mathcal{X}$ and set $Z \subseteq \mathcal{X}$.*

Lemma 6 *If $\neg(X \nrightarrow Y|ZW)$ and $\neg(X \nrightarrow YW|Z)$, 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_{srec} *is complete with respect to $\mathcal{T}_{\mathrm{srec}}$.*

Proof: Let Γ be maximal AX_{srec}–consistent. Let $U,V \in \mathcal{X}$ and maximal $T \subseteq \mathcal{X}$ such that $\neg(U \nrightarrow V|T)$. Let $G \leq G(\Gamma)$ be the $(U,V|T)$–fragment given by Theorem 5. Define the recursive model T_G with singleton $\mathcal{U}_G$ and $\mathcal{D}_X$ equal to the non-negative integers for all $X \in \mathcal{X}$, by[4]

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

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, $T_G \vDash \neg(U \nrightarrow V|Z)$ and $T_G \vDash (\tilde{U} \nrightarrow \tilde{V}|\tilde{T})$ if $(\tilde{U} \nrightarrow \tilde{V}|\tilde{T})$.

Let $\mathcal{U}$ be the collection of all fragments, and T the causal model with domain $\mathcal{U}$ and equations $Y(\cdot; G) = Y^G(\cdot)$ for $G \in \mathcal{U}$. In the rest, we show using induction that T satisfies all literals in Γ; 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 $\neg(X \nrightarrow Y|Z)$, then there is an $(X,Y|Z)$–fragment G such that $X \rightsquigarrow Y$ isn't intercepted by Z. By construction, $T_G \vDash \neg(X \nrightarrow Y|Z)$ which implies $T \vDash \neg(X \nrightarrow Y|Z)$. Assume now $(X \nrightarrow Y|Z)$ and let G be any $(U,V|T)$–fragment. By definition, Z intercepts all paths $X \rightsquigarrow Y$ in G, so by construction $T_G \vDash (X \nrightarrow Y|Z)$. Thus, $T \vDash (X \nrightarrow Y|Z)$.

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

---

[4] An empty max is defined to be 0.

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

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

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

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

We now prove that T_G satisfies $\neg(X_1 \nrightarrow X_4|\emptyset)$ 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, $\neg(X_1 \nrightarrow X_4|X_2)$ which implies $\neg(X_1 \nrightarrow X_4|\emptyset)$. Clearly, the model satisfies $(X_4 \nrightarrow X_3|\emptyset)$, $(X_1 X_3 X_4 \nrightarrow X_2|\emptyset)$ and $(X_2 X_3 X_4 \nrightarrow X_1|\emptyset)$. It also satisfies $(X_1 X_2 \nrightarrow X_4|X_3)$ since the equation for X_4 depends only in X_3. We need to show that it satisfies $(X_1 \nrightarrow X_2 X_4|\emptyset)$;

$$\begin{aligned}[x_1]X_4 &= [x_1, [x_1]X_2]X_4 \\ &= [x_1, X_2 = 0]X_4 \\ &= [x_1, X_2 = 0, [x_1, X_2 = 0]X_3]X_4 \\ &= [x_1, X_2 = 0, X_3 = 0]X_4 \\ &= 0,\end{aligned}$$

independent of x_1. Thus, $T_G \vDash (X_1 \nrightarrow X_2 X_4|\emptyset)$. ∎

3.2 Axiomatization for $\mathcal{T}_{\mathrm{rec}}$

Let Γ be a consistent set with respect to AX_{uniq}. We say that an AX_{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 φ can be derived from set Γ by Recursive Inference (RI) if and only if φ holds in all recursive extensions of Γ.*

As with SRI, all RI inferences are decidable. Let AX_{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 $\mathcal{T}_{\mathrm{rec}}$.

Theorem 8 AX_{rec} *is sound with respect to $\mathcal{T}_{\mathrm{rec}}$.*

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

$(X \nrightarrow Y|Z)$, then the empty graph is a $(X,Y|Z)$–fragment. If $\neg(X \nrightarrow Y|Z)$, then there is $u \in \mathcal{U}$ such that $T(u) \vDash \neg(X \nrightarrow Y|Z)$. By Theorem 5, $G(\text{Th}\, T(u))$ contains a $(X,Y|Z)$–fragment. Then, $\text{Th}\, T$ is a recursive extension of Γ. ∎

Let T_1 and T_2 be two functional causal models with exogenous domains $\mathcal{U}_1, \mathcal{U}_2$, and equations F_Y^i for $i = 1, 2$ and $Y \in \mathcal{X}$. 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 $\mathcal{U} = \{1\} \times \mathcal{U}_1 \cup \{2\} \times \mathcal{U}_2$ and equations given by $F_Y(\cdot\,; (i,u)) \overset{\text{def}}{=} F_Y^i(\cdot\,; u)$ 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 $\mathcal{U} = \bigcup_i \{i\} \times \mathcal{U}_i$ and equations given by $F_Y(\cdot\,; (i,u)) = F_Y^i(\cdot\,; u)$ for $1 \leq i \leq n$.

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

(i) $T \cong T_1 \oplus \cdots \oplus T_n$ *if and only if* $\text{Th}\, T$ *is the unique* AX_{uniq}*–extension of*

$$\bigcup_{i=1}^{n} \text{Lit}^- (\text{Th}\, T_i) \;\; \cup \;\; \bigcap_{i=1}^{n} \text{Lit}^+ (\text{Th}\, T_i) \,. \quad (3)$$

(ii) *If* $T \cong T1 \oplus \cdots \oplus T_n$, *then* $G(T) = G(T_1) \cup \cdots \cup G(T_n)$.

Proof: *(i).* Denote with Γ the set of formulas in Eq.(3). First we prove that Γ has only one extension. Let $\varphi \in \mathcal{L}$ be a positive literal. If $\exists i[T_i \nvDash \varphi]$, then $\exists i[T_i \vDash \neg\varphi]$ which implies $\neg\varphi \in \Gamma$. If $\forall i[T_i \vDash \varphi]$, then $\varphi \in \cap_{i=1}^{n}\text{Lit}^+ (\text{Th}\, T_i)$ which implies $\varphi \in \Gamma$. For negative literals consider $\psi = \neg\varphi$. Thus, Γ is a maximal consistent set, so its unique extension is $\text{Th}_{\text{uniq}}\, \Gamma$.

Assume $T \cong T_1 \oplus \cdots \oplus T_n$, and φ a positive literal. If $T \vDash \varphi$, then $\forall(i,u)[T(i,u) \vDash \varphi]$ which implies $\forall i[\varphi \in \text{Lit}^+ (\text{Th}\, T_i)]$, so $\varphi \in \Gamma$. If $T \nvDash \varphi$, then $\exists(i,u)[T(i,u) \nvDash \varphi]$ which implies $\exists i[\varphi \notin \text{Lit}^+ (\text{Th}\, T_i)]$, so $\varphi \notin \Gamma$. Therefore, $\text{Lit}\,(\text{Th}\, T) = \text{Lit}\, \Gamma$ which implies the result.

Now, assume $\text{Th}\, T = \text{Th}_{\text{uniq}}\, \Gamma$. Let φ be a positive literal. Then, $\varphi \in \text{Th}\, T$ iff

$$\varphi \in \bigcap_{i=1}^{n} \text{Lit}^+ (\text{Th}\, T_i) \iff \varphi \in \text{Th}\, (T_1 \oplus \cdots \oplus T_n).$$

(ii). Assume $T \cong T_1 \oplus \cdots \oplus T_n$. Let X, Y be two variables and $W = (XY)^c$. Then, $(X,Y) \in G(T)$ iff $(X,Y) \in G(\text{Th}\,(T_1 \oplus \cdots \oplus T_n))$ iff $\exists i\big[T_i \vDash \neg(X \nrightarrow Y|W)\big]$ iff $\exists i\big[(X,Y) \in G(T_i)\big]$. ∎

Completeness

To motivate our proof for completeness, observe that a model $T \in \mathcal{T}_{\text{rec}}$ is composed of different $\mathcal{T}_{\text{srec}}$ 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 Γ and negative literal $\varphi \in \Gamma$, we assign a maximal AX_{srec}–consistent set $\Gamma' \subseteq \Gamma$ containing φ. All such sets are then use to build $\mathcal{T}_{\text{srec}}$ models that combine into a model $T \in \mathcal{T}_{\text{rec}}$ such that $T \vDash \Gamma$.

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

$$\Psi_0 = \text{Th}_{\text{srec}} \left(\{\varphi\} \cup \text{Lit}^+ \Gamma\right)$$

$$\Psi_{i+1} = \begin{cases} \text{Th}_{\text{srec}} \left(\Psi_i \cup \{\psi_i\}\right) & \text{if } \Psi_i \cup \{\psi_i\} \text{ is} \\ & \text{AX}_{\text{srec}}\text{–consistent,} \\ \text{Th}_{\text{srec}} \left(\Psi_i \cup \{\neg\psi_i\}\right) & \text{otherwise.} \end{cases}$$

We want to prove that each Ψ_i is AX_{srec}–consistent. By construction, it is enough to show that Ψ_0 is consistent. Fact that is true since Γ is AX_{rec}–consistent; i.e., $G(\Gamma)$ contains an $(X,Y|Z)$–fragment for φ. Note that Ψ_n is maximal and that $\text{Lit}^- \Psi_i \subseteq \text{Lit}^- \Gamma$ since $\text{Lit}^+ \Gamma \subseteq \Psi_0$. We call Ψ_n a (φ, Γ)–*foliation*.

Theorem 10 AX_{rec} *is complete with respect to* $\mathcal{T}_{\text{rec}}$.

Proof: Let Γ be a maximal AX_{rec}–consistent, and Ψ a (φ, Γ)–foliation. Denote with T_Ψ the strong recursive model such that $T_\Psi \vDash \Psi$ (it exists by completeness of AX_{srec}). 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 $\neg(U \nrightarrow Z|T)$ in Γ, there exists a foliation Ψ such that $T_\Psi \vDash \neg(U \nrightarrow Z|T)$. Let $\mathcal{F} = \{\Psi : \Psi \text{ is a } (\varphi, \Gamma)$–foliation $\}$, and $T = \bigoplus_{\Psi \in \mathcal{F}} T_\Psi$. Then, by Theorem 9, $T \vDash \Gamma$. ∎

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

Theorem 11 $\mathcal{T}_{\text{rec}} < \mathcal{T}_{\text{uniq}}$.

Proof: Consider the model T with variables $\{X_1, \ldots, X_4\}$, singleton $\mathcal{U}$ and equations

$$X_1 = 1, \quad X_2 = (X_1, X_3), \quad X_3 = X_2\!\downarrow_1, \quad X_4 = X_2\!\downarrow_2 \,.$$

It is a non-recursive model that satisfies $\neg(X_1 \nrightarrow X_4|\emptyset)$ and $(X_1 \nrightarrow X_4|Z)$ for $Z \in \{X_2, X_3\}$. Note that $G(T) = G(\text{Th}\, T)$ is the graph with edges $\{X_1 \to X_2 \to X_4, X_2 \rightleftarrows X_3\}$, and that T satisfies $\neg(X_2 \nrightarrow X_4|X_1X_3)$, $\neg(X_1 \nrightarrow X_3|X_4)$ and $\neg(X_2 \nrightarrow X_3|X_4)$. It can be shown that $\text{Th}\, T$ has no recursive extension, so it is not AX_{rec}–consistent. Therefore, there is no recursive model $\mathcal{L}$–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 Γ be a set of causal relevances that hold in T; i.e., $T \vDash \Gamma$. If $T \in \mathcal{T}_{\mathrm{srec}}$, then the completeness of $\mathrm{AX}_{\mathrm{srec}}$ and Theorem 3 imply $G(T)$ is equal to $G(\Gamma')$ for some $\mathrm{AX}_{\mathrm{srec}}$–extension Γ' of Γ. Thus, the graph defined as

$$\mathcal{G}(\Gamma) \stackrel{\mathrm{def}}{=} \bigcap_{\Gamma'} G(\Gamma'),$$

where the intersection is over all $\mathrm{AX}_{\mathrm{srec}}$–extensions of Γ, is an edge subgraph of $G(T)$. Completeness is needed here in order to guarantee that every Γ' has a corresponding model in $\mathcal{T}_{\mathrm{srec}}$. 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 Γ_i. A rational approach would be to buy the set that minimizes the pondered cost $C_i \stackrel{\mathrm{def}}{=} c_i / \#(\mathcal{G}(\Gamma_i) - \mathcal{G}(\Gamma))$ 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 Γ_i generates no new edges, then $C_i = \infty$. Similar definitions can be made when $T \in \mathcal{T}_{\mathrm{rec}}$. 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 $\neg(X \nrightarrow Y | Z)$, then the graph we are looking for must has a path $X \rightsquigarrow Y$ not intercepted by Z.

Another application of the calculus is for testing whether a model is recursive or not. Since $\mathrm{AX}_{\mathrm{rec}}$ is complete with respect to $\mathcal{T}_{\mathrm{rec}}$, then a *necessary* test for recursiveness is $\mathrm{AX}_{\mathrm{rec}}$–consistency. Thus, if Γ is not $\mathrm{AX}_{\mathrm{rec}}$–consistent, then Γ 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 $\mathcal{T}_{\mathrm{uniq}}$ into the two subclasses $\mathcal{T}_{\mathrm{srec}}$ and $\mathcal{T}_{\mathrm{rec}}$. This decomposition is necessary for the formalization of causal relevance from the three-place relation $(X \nrightarrow Y | Z)$, and it is useful as shown in the applications of the calculus. The second contribution is the collection of axiomatizations $\mathrm{AX}_{\mathrm{uniq}}$, $\mathrm{AX}_{\mathrm{rec}}$ and $\mathrm{AX}_{\mathrm{srec}}$. $\mathrm{AX}_{\mathrm{uniq}}$ extends the system in [6] with the three new axioms A5, A8 and A9. It is a sound axiomatization with respect to the class $\mathcal{T}_{\mathrm{uniq}}$. $\mathrm{AX}_{\mathrm{rec}}$ is a stronger version of $\mathrm{AX}_{\mathrm{uniq}}$ that is sound and complete with respect to $\mathcal{T}_{\mathrm{rec}}$. It is equal to $\mathrm{AX}_{\mathrm{uniq}}$ plus the rule RI. The axiomatization $\mathrm{AX}_{\mathrm{srec}}$,

a stronger version $\mathrm{AX}_{\mathrm{rec}}$, is sound and complete with respect to $\mathcal{T}_{\mathrm{srec}}$. It is equal to $\mathrm{AX}_{\mathrm{uniq}}$ 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 $\mathrm{AX}_{\mathrm{uniq}}$ is a complete system with respect to $\mathcal{T}_{\mathrm{uniq}}$, if the collection A1–A9 is an independent set of Axioms, and which Axioms can be used to replace the RI and SRI rules.

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.