In Proceedings of the Twenty-First National Conference on Artificial Intelligence, AAAI Press, Menlo Park, CA, 1219-1226, July 2006.

TECHNICAL REPORT
R-327
February 2006

Identification of Joint Interventional Distributions in Recursive Semi-Markovian Causal Models

Ilya Shpitser, Judea Pearl
Cognitive Systems Laboratory
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA. 90095
{ilyas, judea}@cs.ucla.edu

Abstract

This paper is concerned with estimating the effects of actions from causal assumptions, represented concisely as a directed graph, and statistical knowledge, given as a probability distribution. We provide a necessary and sufficient graphical condition for the cases when the causal effect of an arbitrary set of variables on another arbitrary set can be determined uniquely from the available information, as well as an algorithm which computes the effect whenever this condition holds. Furthermore, we use our results to prove completeness of do-calculus [Pearl, 1995], and a version of an identification algorithm in [Tian, 2002] for the same identification problem.

Introduction

This paper deals with computing effects of actions in domains specified as $causal\ diagrams$, or graphs with directed and bidirected edges. Vertices in such graphs correspond to variables of interest, directed edges correspond to potential direct causal relationships between variables, and bidirected edges correspond to 'hidden common causes,' or spurious dependencies between variables [Pearl, 1995], [Pearl, 2000]. Aside from causal knowledge encoded by these graphs, we also have statistical knowledge in the form of a joint probability distribution over observable variables, which we will denote by P.

An action on a variable set $\mathbf{X}$ in a causal domain consists of forcing $\mathbf{X}$ to particular values $\mathbf{x}$, without regard for the normal causal relationships that determine the values of $\mathbf{X}$. This action, denoted $do(\mathbf{x})$ in [Pearl, 2000], changes the original joint distribution P over observables into a new $interventional\ distribution$ denoted $P_{\mathbf{x}}$. The marginal distribution $P_{\mathbf{x}}(\mathbf{Y})$ of a variable set $\mathbf{Y}$ obtained from $P_{\mathbf{x}}$ will be our notion of effect of action $do(\mathbf{x})$ on $\mathbf{Y}$.

Our task is to characterize cases when $P_{\mathbf{x}}(\mathbf{Y})$ can be determined uniquely from P, or $identified$ in a given graph G. It is well known that in Markovian models, those causal domains whose graphs do not contain bidirected edges, all effects are identifiable [Pearl, 2000]. If our model contains 'hidden common causes,' that is if the model is semi-Markovian, the situation is less clear.

Multiple sufficient conditions for identifiability in the semi-Markovian case are known [Spirtes, Glymour, & Scheines, 1993], [Pearl & Robins, 1995], [Pearl, 1995], [Kuroki & Miyakawa, 1999]. [Pearl, 2000] contains an excellent summary of these results. These results generally take advantage of the fact that certain properties of the causal diagram reflect properties of P. These results are thus phrased in the language of graph theory. For example, the back-door criterion [Pearl, 2000], states that if there exists a set $\mathbf{Z}$ of non-descendants of $\mathbf{X}$ that 'block' certain paths in the graph from $\mathbf{X}$ to $\mathbf{Y}$, then $P_{\mathbf{x}}(\mathbf{Y}) = \sum_{\mathbf{z}} P(\mathbf{Y}|\mathbf{z}, \mathbf{x})P(\mathbf{z})$.

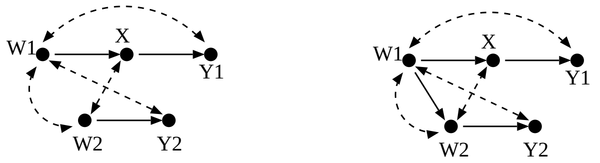
Results in [Pearl, 1995], [Halpern, 2000] take a different approach, and provide sound rules which are used to manipulate the expression corresponding to the effect algebraically. These rules are then applied until the resulting expression can be computed from P.

Though the axioms in [Halpern, 2000] were shown to be complete, the practical applicability of the result is limited, since it does not provide a closed form criterion for the cases when effects are not identifiable, nor a closed form algorithm for expressing effects in terms of P when they are identifiable. Instead, one must rely on finding a good proof strategy and hope the effect expression is reduced to something derivable from P.

Recently, a number of necessity results for identifiability have been proven. One such result [Tian & Pearl, 2002] states that P_x is identifiable if and only if there is no path consisting entirely of bidirected arcs from X to a child of X.

The authors have also been made aware of a paper currently in review [Huang & Valtorta, 2006] which shows a modified version of an algorithm found in [Tian, 2002] is complete for identifying $P_{\mathbf{x}}(\mathbf{y})$, where $\mathbf{X}$, $\mathbf{Y}$ are sets. The result in this paper uses non-positive distributions, but has a simpler proof, and was derived independently.

In this paper we close the problem of identification of causal effects of the form $P_{\mathbf{x}}(\mathbf{y})$ in semi-Markovian causal models. In particular, we show that anytime such an effect is non-identifiable there exists a corresponding graphical structure which we call a $hedge$. We use this graphical criterion to construct a sound and complete algorithm for identification of $P_{\mathbf{x}}(\mathbf{y})$ from P. The algorithm returns either an expression derivable from P or a hedge which witnesses the non-identifiability of the effect. We also show that steps of our algorithm correspond to sequences of applications of rules of do-calculus [Pearl, 1995], thus proving the completeness of do-calculus for the same identification problem. Furthermore, we show a version of a known identification algorithm [Tian, 2002] is also complete and thus equivalent to our algorithm.



(a) (b)

Figure 1: (a) Graph G. (b) Graph G'.
W_1 - dating habits, X - use of contraception, Y_1 - incidence of STDs, W_2 - alcohol use, and Y_2 - traffic fatalities.

An Example of Causal Modeling

In this section, we illustrate the identifiability problem with an example. Assume we are studying lifestyle-related causes of early death in young adults. For our example, we restrict our attention to deaths in traffic accidents, as well as deadly sexually transmitted diseases like hepatitis B, and AIDS. A simple causal model listing some causes of such deaths is shown in Fig. 1 (a). In this model, a person's dating habits influence their likelihood of using contraception, which in turn influences the likelihood of contracting a deadly STD. There is also a confounding variable between STDs and dating habits, perhaps related to the kind of social circle a given person is in. Similarly, alcohol use influences the likelihood of dying in a car accident. At the same time, alcohol consumption might be related to the use of contraceptives via a confounder representing moral or religious convictions. Alcohol and dating habits might be linked by personality, and dating habits and traffic fatalities might be linked by the person's financial situation (a person might not be able to afford a car).

We also consider a related situation shown in Fig. 1 (b), where dating habits have a direct influence on alcohol use – it could be argued that 'bar flies' are more likely to consume alcohol regardless of their personality.

Assuming this model is an accurate representation of the situation we wish to study, we might be interested in what effect a lifestyle change, say a consistent use of contraceptives, has on causes of death under consideration. We represent said lifestyle change as an action on the variable X, which disregards the normal functional relationships imposed by the model, thus severing X from the influence of its parents and fixing its value. We are now interested in the distribution this action induces on the variables Y_1, and Y_2, representing the causes of death. Since it is not possible to impose such changes on people in an experimental setting, we wish to identify the interventional distribution $P_x(Y_1, Y_2)$ from the distribution we can obtain without forcing anyone, namely $P(X, W_1, W_2, Y_1, Y_2)$. Characterizing the situations where latter, observational distributions can be used to infer the former is the causal effect identification problem.

The subtlety of this problem can be illustrated by noting that in the graph in Fig. 1 (a) the effect in question is identifiable from $P(\mathbf{V})$, whereas in the graph in Fig. 1 (b) it is not. The subsequent sections will shed some light on why this is so.

Notation and Definitions

In this section we reproduce the technical definitions needed for the rest of the paper, and introduce common non-identifying graph structures. We will denote variables by capital letters, and their values by small letters. Similarly, sets of variables will be denoted by bold capital letters, sets of values by bold small letters. We will also use some graph-theoretic abbreviations: $Pa(\mathbf{Y})_G$, $An(\mathbf{Y})_G$, and $De(\mathbf{Y})_G$ will denote the set of (observable) parents, ancestors, and descendants of the node set $\mathbf{Y}$ in G, respectively. We will omit the graph subscript if the graph in question is assumed or obvious. We will denote the set $\{X \in G | De(X)_G = \emptyset\}$ as the root set of G. A graph $G_{\mathbf{Y}}$ will denote a subgraph of G containing nodes in $\mathbf{Y}$ and all arrows between such nodes. Finally, following [Pearl, 2000], $G_{\overline{\mathbf{x}}\underline{\mathbf{y}}}$ stands for the edge subgraph of G where all in coming arrows into $\mathbf{X}$ and all outgoing arrows from $\mathbf{Y}$ are deleted.

Having fixed our notation, we can proceed to formalize the notions discussed in the previous section.

Definition 1 *A probabilistic causal model (PCM) is a tuple* $M = \langle \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{F}, P(\boldsymbol{U}) \rangle$, *where*

(i) $\boldsymbol{U}$ *is a set of background or exogenous variables, which cannot be observed or experimented on, but which can influence the rest of the model.*

(ii) $\boldsymbol{V}$ *is a set* $\{V_1, ..., V_n\}$ *of observable or endogenous variables. These variables are considered to be functionally dependent on some subset of* $\boldsymbol{U} \cup \boldsymbol{V}$.

(iii) $\boldsymbol{F}$ *is a set of functions* $\{f_1, ..., f_n\}$ *such that each* f_i *is a mapping from a subset of* $\boldsymbol{U} \cup \boldsymbol{V} \setminus \{V_i\}$ *to* V_i, *and such that* $\bigcup \boldsymbol{F}$ *is a function from* $\boldsymbol{U}$ *to* $\boldsymbol{V}$.

(iv) $P(\boldsymbol{U})$ *is a joint probability distribution over the variables in* $\boldsymbol{U}$.

For a variable V, its corresponding function is denoted as f_V. For the purposes of this paper, we assume all variable domains are finite. The distribution on $\mathbf{V}$ induced by $P(\mathbf{U})$ and $\mathbf{F}$ will be denoted $P(\mathbf{V})$.

Sometimes it is assumed $P(\mathbf{V})$ is a positive distribution. In this paper we do not make this assumption. Thus, we must make sure that for every distribution $P(\mathbf{W}|\mathbf{Z})$ that we consider, $P(\mathbf{Z})$ must be positive. This can be achieved by making sure to sum over events with positive probability only. Furthermore, for any action $do(\mathbf{x})$ that we consider, it must be the case that $P(\mathbf{x}|Pa(\mathbf{X})_G \setminus \mathbf{X}) > 0$ otherwise the distribution $P_{\mathbf{x}}(\mathbf{V})$ is not well defined [Pearl, 2000].

The induced graph G of a causal model M contains a node for every element in $\mathbf{V}$, a directed edge between nodes X and Y if f_Y possibly uses the values of X directly to determine the value of Y, and a bidirected edge between nodes X and Y if f_X and f_Y both possibly use the value of some variable in $\mathbf{U}$ to determine their respective values. In this paper we consider $recursive$ causal models, those models which induce acyclic graphs.

In the framework of causal models, actions are modifications of functional relationships. Each action $do(\mathbf{x})$ on a causal model

M produces a new model, with a new probability distribution. Formally, this is defined as follows.

Definition 2 (submodel) *For a causal model $M = \langle \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{F}, P(\boldsymbol{U}) \rangle$, an intervention $do(\boldsymbol{x})$ produces a new causal model $M_{\boldsymbol{x}} = \langle \boldsymbol{U}, \boldsymbol{V_x}, \boldsymbol{F_x}, P(\boldsymbol{U}) \rangle$, where $\boldsymbol{V_x}$ is a set of distinct copies of variables in $\boldsymbol{V}$, and $\boldsymbol{F_x}$ is obtained by taking distinct copies of functions in $\boldsymbol{F}$, but replacing all copies of functions which determine the variables in $\boldsymbol{X}$ by constant functions setting the variables to values $\boldsymbol{x}$. The model $M_{\boldsymbol{x}}$ is called the submodel of M, and the distribution over $\boldsymbol{V_x}$ is denoted $P_{\boldsymbol{x}}(\boldsymbol{V})$.*

Since subscripts are used to denote submodels, we will use numeric superscripts to enumerate models (e.g. M^1). For a model M^i, we will often denote it's associated probability distributions as P^i rather than P.

We can now define formally the notion of identifiability of interventions from observational distributions.

Definition 3 (Causal Effect Identifiability) *The causal effect of an action $do(\boldsymbol{x})$ on a set of variables $\boldsymbol{Y}$ such that $\boldsymbol{Y} \cap \boldsymbol{X} = \emptyset$ is said to be identifiable from P in G if $P_{\boldsymbol{x}}(\boldsymbol{Y})$ is (uniquely) computable from $P(\boldsymbol{V})$ in any causal model which induces G.*

The following lemma establishes the conventional technique used to prove non-identifiability in a given G.

Lemma 1 *Let $\boldsymbol{X}, \boldsymbol{Y}$ be two sets of variables. Assume there exist two causal models M^1 and M^2 with the same induced graph G such that $P^1(\boldsymbol{V}) = P^2(\boldsymbol{V})$, $P^1(\boldsymbol{x}|Pa(\boldsymbol{X})_G \setminus \boldsymbol{X}) > 0$, and $P^1_{\boldsymbol{x}}(\boldsymbol{Y}) \neq P^2_{\boldsymbol{x}}(\boldsymbol{Y})$. Then $P_{\boldsymbol{x}}(\boldsymbol{y})$ is not identifiable in G.*

Proof: No function from P to $P_{\mathbf{x}}(\mathbf{y})$ can exist by assumption, let alone a computable function. □

The simplest example of a non-identifiable graph structure is the so called 'bow arc' graph. See Fig. 2 (a). The bow arc graph has two endogeous nodes: X, and its child Y. Furthermore, X and Y have a hidden exogenous parent U. Although it is well known that $P_x(Y)$ is not identifiable in this graph, we give a simple proof here which will serve as a seed of a similar proof for more general graph structures.

Theorem 1 *$P_x(Y)$ is not identifiable in the bow arc graph.*

Proof: We construct two causal models M^1 and M^2 such that $P^1(X,Y) = P^2(X,Y)$, and $P^1_x(Y) \neq P^2_x(Y)$. The two models agree on the following: all 3 variables are boolean, U is a fair coin, and $f_X(u) = u$. Let $\oplus$ denote the exclusive or (XOR) function. Then Y is equal to $u \oplus x$ in M^1 and Y is equal to 0 in M^2. It's not difficult to see that $P^1(X,Y) = P^2(X,Y)$, while $P^2_x(Y) \neq P^1_x(Y)$. Note that while P is non-positive, it is straightforward to modify the proof for the positive case by letting f_Y functions in both models return 1 half the time, and the values outlined above half the time. □

A number of other specific graphs have been shown to contain unidentifiable effects. For instance, in all graphs in Fig. 2, taken from [Pearl, 2000], $P_x(Y)$ is not identifiable.

C-Trees and Direct Effects

A number of example graphs from Fig. 2 have the following two properties. Firstly, Y is the bottom-most node of the graph. Secondly, all the nodes in the graph are pairwise connected by bidirected paths. Sets of nodes interconnected by bidirected paths turned out to be an important notion for identifiability and have been studied at length in [Tian, 2002] under the name of C-$components$.

Definition 4 (C-component) *Let G be a semi-Markovian graph such that a subset of its bidirected arcs forms a spanning tree over all vertices in G. Then G is a C-component (confounded component).*

If G is not a C-component, it can be uniquely partitioned into a set $C(G)$ of subgraphs, each a maximal C-component.

An important result states that for any set **C** which is a C-component, in a causal model M with graph G, $P_{\mathbf{V} \setminus \mathbf{C}}(\mathbf{C})$ is identifiable [Tian, 2002]. The quantity $P_{\mathbf{V} \setminus \mathbf{C}}(\mathbf{C})$ will also be denoted as $Q[\mathbf{C}]$. For the purposes of this paper, C-components are important because a distribution P in a semi-Markovian graph G factorizes such that each product term $Q[\mathbf{C}]$ corresponds to a C-component. For instance, the graphs shown in Fig. 2 (b) and (c), both have 2 C-components: $\{X, Z\}$ and $\{Y\}$. Thus, the corresponding distribution factorizes as $P(x, z, y) = P_y(x, z)P_{x,z}(y)$. It is this factorization which will ultimately allow us to decompose the identification problem into smaller subproblems, and thus construct an identification algorithm.

Noting that many graphs from the previous section with non-identifiable effects $P_x(Y)$ were both C-components and ancestral sets of Y, we consider the following structure.

Definition 5 (C-tree) *Let G be a semi-Markovian graph such that $C(G) = \{G\}$, all observable nodes have at most one child, and there is a node Y, which is a descendant of all nodes. Then G is a Y-rooted C-tree (confounded tree).*

The graphs in Fig. 2 (a) (d) (e) (f) and (h), including the bow-arc graph, are Y-rooted C-trees.

There is a relationship between C-trees and interventional distributions of the form $P_{pa(Y)}(Y)$. Such distributions are known as $direct\ effects$, and correspond to the influence of a variable X on its child Y along some edge, where the variables $Pa(Y) \setminus \{X\}$ are fixed to some reference values.

Direct effects are of great importance in the legal domain, where one is often concerned with whether a given party was directly responsible for damages, as well as medicine, where elucidating the direct effect of medication, or disease on the human body in a given context is crucial. See [Pearl, 2000],[Pearl, 2001] for a more complete discussion of direct effects.

We now show that the absence of Y-rooted C-trees in G means the direct effect on Y is identifiable.

Theorem 2 *Let M be a causal model with graph G. Then for any node Y, the direct effect $P_{pa(Y)}(Y)$ is identifiable if there is no subgraph of G which forms a Y-rooted C-tree.*

Proof: From [Tian, 2002], we know that whenever there is no subgraph G' of G, such that all nodes in G' are ancestors of Y, and G' is a C-component, $P_{pa(Y)}(Y)$ is identifiable. Clearly, if no such G' exists, then no Y-rooted C-tree exists. On the other hand, if G' does exist, then it is possible to remove a set of edges from G' to obtain a Y-rooted C-tree. □

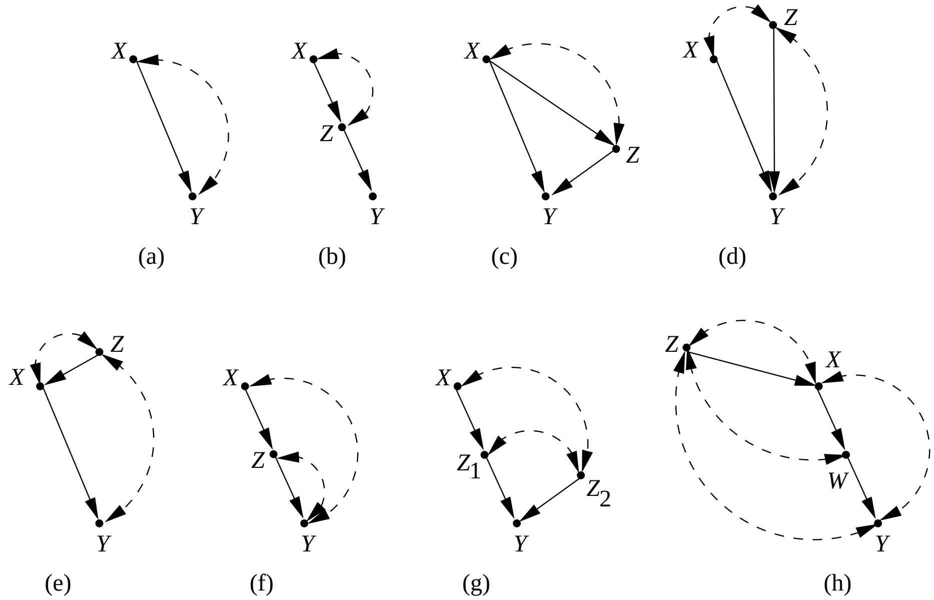


Figure 2: Unidentifying graphs.

Theorem 2 suggests that C-trees are troublesome structures for the purposes of identification of direct effects. In fact, our investigation revealed that Y-rooted C-trees are troublesome for any effect on Y, as the following theorem shows.

Theorem 3 *Let G be a Y-rooted C-tree. Then the effects of any set of nodes in G on Y are not identifiable.*

Proof: See Appendix. □

While this theorem closes the identification problem for direct effects, the problem of identifying general effects on a single variable Y is more subtle, as the following corollary shows.

Corollary 1 *Let G be a semi-Markovian graph, let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets of variables. If there exists a node W which is an ancestor of some node $Y \in \boldsymbol{Y}$ and such that there exists a W-rooted C-tree which contains any variables in $\boldsymbol{X}$, then $P_{\mathbf{x}}(\boldsymbol{Y})$ is not identifiable.*

Proof: Fix a W-rooted C-tree T, and a path p from W to Y. Consider the graph $p \cup T$. Note that in this graph $P_{\mathbf{x}}(Y) = \sum_w P_{\mathbf{x}}(w)P(Y|w)$. It is now easy to construct $P(Y|W)$ in such a way that the mapping from $P_{\mathbf{x}}(W)$ to $P_{\mathbf{x}}(Y)$ is one to one, while making sure P is positive. □

This corollary implies that the effect of $do(\mathbf{x})$ on a given singleton Y can be non-identifiable even if Y is nowhere near a C-tree, as long as the effect of $do(\mathbf{x})$ on a set of ancestors of Y is non-identifiable. Therefore identifying effects on a single variable is not really any easier than the general problem of identifying effects on multiple variables. We consider this general problem in the next section.

Finally, we note that the last two results relied on existence of a C-tree without giving an explicit algorithm for constructing one. In the remainder of the paper we will give an algorithm which, among other things, will construct the necessary C-tree, if it exists.

C-Forests, Hedges, and Non-Identifiable Effects

The previous section establishes a powerful necessary condition for the identification of effects on a single variable. It is the natural next step to ask whether a similar condition exists for effects on multiple variables. We start by considering a multi-vertex generalization of a C-tree.

Definition 6 (C-forest) *Let G be a semi-Markovian graph, where $\boldsymbol{Y}$ is the maximal root set. Then G is a $\boldsymbol{Y}$-rooted $C\text{-}forest$ if G is a C-component, and all observable nodes have at most one child.*

We will show that just as there is a close relationship between C-trees and direct effects, there is a close relationship between C-forests and general effects of the form $P_{\mathbf{x}}(\mathbf{Y})$, where $\mathbf{X}$ and $\mathbf{Y}$ are sets of variables.

Definition 7 (hedge) *Let $\boldsymbol{X}, \boldsymbol{Y}$ be sets of variables in G. Let F, F' be $\boldsymbol{R}$-rooted C-forests such that $F \cap \boldsymbol{X} \neq \emptyset$, $F' \cap \boldsymbol{X} = \emptyset$, $F' \subseteq F$, and $\boldsymbol{R} \subset An(\boldsymbol{Y})_{G_{\overline{\mathbf{x}}}}$. Then F and F' form a hedge for $P_{\mathbf{x}}(\boldsymbol{y})$ in G.*

The mental picture for a hedge is as follows. We start with a C-forest F'. Then, F' 'grows' new branches, while retaining the same root set, and becomes F. Finally, we 'trim the hedge,' by performing the action $do(\mathbf{x})$ which has the effect of removing some incoming arrows in $F \setminus F'$. It's easy to check that every graph in Fig. 2 contains a pair of C-forests that form a hedge for $P_x(Y)$. Similarly, the graph in Fig. 1 (a) does not contain C-forests forming a hedge for $P_x(Y_1, Y_2)$, while the graph in Fig. 1 (b) does: if e is the edge between W_1 and X, then $F = G \setminus \{e\}$, and $F' = F \setminus \{X\}$. Note that for the special case of C-trees, F is the C-tree itself, and F' is the singleton root Y.

Theorem 4 (hedge criterion) *Assume there exist C-forests F, F' that form a hedge for $P_{\mathbf{x}}(\boldsymbol{y})$ in G. Then $P_{\mathbf{x}}(\boldsymbol{y})$ is not identifiable in G.*

function **ID**($\mathbf{y}$, $\mathbf{x}$, P, G)
INPUT: $\mathbf{x}$,$\mathbf{y}$ value assignments, P a probability distribution, G a causal diagram.
OUTPUT: Expression for $P_{\mathbf{x}}(\mathbf{y})$ in terms of P or **FAIL**(F,F').

1 if $\mathbf{x} = \emptyset$ return $\sum_{\mathbf{v} \setminus \mathbf{y}} P(\mathbf{v})$.

2 if $\mathbf{V} \setminus An(\mathbf{Y})_G \neq \emptyset$
return $\mathbf{ID}(\mathbf{y}, \mathbf{x} \cap An(\mathbf{Y})_G, \sum_{\mathbf{v} \setminus An(\mathbf{Y})_G} P, An(\mathbf{Y})_G)$.

3 let $\mathbf{W} = (\mathbf{V} \setminus \mathbf{X}) \setminus An(\mathbf{Y})_{G_{\overline{\mathbf{x}}}}$.
if $\mathbf{W} \neq \emptyset$, return $\mathbf{ID}(\mathbf{y}, \mathbf{x} \cup \mathbf{w}, P, G)$.

4 if $C(G \setminus \mathbf{X}) = \{S_1, ..., S_k\}$
return $\sum_{\mathbf{v} \setminus (\mathbf{y} \cup \mathbf{x})} \prod_i \mathbf{ID}(s_i, \mathbf{v} \setminus s_i, P, G)$.

if $C(G \setminus \mathbf{X}) = \{S\}$

5 if $C(G) = \{G\}$, throw $\mathbf{FAIL}(G, S)$.

6 if $S \in C(G)$ return $\sum_{s \setminus \mathbf{y}} \prod_{\{i | V_i \in S\}} P(v_i | v_G^{(i-1)})$.

7 if $(\exists S') S \subset S' \in C(G)$ return $\mathbf{ID}(\mathbf{y}, \mathbf{x} \cap S',$
$\prod_{\{i | V_i \in S'\}} P(V_i | V_G^{(i-1)} \cap S', v_G^{(i-1)} \setminus S'), S')$.

Figure 3: An improved identification algorithm. **FAIL** propagates through recursive calls like an exception, and returns F, F' which form the hedge which witnesses non-identifiability.

Proof: See Appendix. □

Hedges generalize the complete condition for identification of P_x from P in [Tian & Pearl, 2002] which states that if Y is a child of X and there exists a set of bidirected arcs connecting X to Y then (and only then) P_x is not identifiable. Let G consist of X, Y and the nodes $W_1, ..., W_k$ on the bidirected path from X to Y. It is not difficult to check that G forms a hedge for $P_x(Y, W_1, ..., W_k)$. Since hedges characterize a wide variety of non-identifiable graphs, and generalize a known complete condition for a special case of the identification problem, it might be reasonable to suppose that the hedge criterion is a complete characterization of non-identifiability in graphical causal models. To prove this supposition, we would need to construct an algorithm which identifies any effect lacking a hedge. This algorithm is the subject of the next section.

A Complete Identification Algorithm

Given the characterization of unidentifiable effects in the previous section, we can attempt to solve the identification problem in all other cases, and hope for completeness. To do this we construct an algorithm that systematically takes advantage of the properties of C-components to decompose the identification problem into smaller subproblems until either the entire expression is identified, or we run into the problematic hedge structure. This algorithm, called **ID**, is shown in Fig. 3.

Before showing the soundness and completeness properties of **ID**, we give the following example of the operation of the algorithm. Consider the graph G in Fig. 1 (a), where we want to identify $P_x(y_1, y_2)$ from $P(\mathbf{V})$.

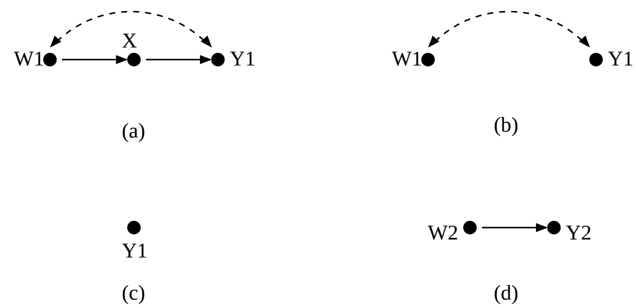


Figure 4: Subgraphs of G used for identifying $P_x(y_1, y_2)$.

Since $G = An(\{Y_1, Y_2\})_G, C(G \setminus \{X\}) = \{G\}$, and $\mathbf{W} = \{W_1\}$, we invoke line 3 and attempt to identify $P_{x,w}(y_1, y_2)$. Now $C(G \setminus \{X, W\}) = \{Y_1, W_2 \rightarrow Y_2\}$, so we invoke line 4. Thus the original problem reduces to identifying $\sum_{w_2} P_{x,w_1,w_2,y_2}(y_1) P_{w,x,y_1}(w_2, y_2)$.

Solving for the second expression, we trigger line 2, noting that we can ignore nodes which are not ancestors of W_2 and Y_2, which means $P_{w,x,y_1}(w_2, y_2) = P(w_2, y_2)$. The corresponding G is shown in Fig. 4 (d).

Solving for the first expression, we first trigger line 2 also, obtaining $P_{x,w_1,w_2,y_2}(y_1) = P_{x,w}(y_1)$. The corresponding G is shown in Fig. 4 (a). Next, we trigger line 7, reducing the problem to computing $P_w(y_1)$ from $P(Y_1|X, W_1)P(W_1)$. The corresponding G is shown in Fig. 4 (b). Finally, we trigger line 2, obtaining $P_w(y_1) = \sum_{w_1} P(y_1|x, w_1)P(w_1)$. The corresponding G is shown in Fig. 4 (c). Putting everything together, we obtain: $P_x(y_1, y_2) = \sum_{w_2} P(y_1, w_2) \sum_{w_1} P(y_1|x, w_1)P(w_1)$.

As we showed before, the very same effect $P_x(y_1, y_2)$ in a very similar graph G' shown in Fig. 1 (b) is not identifiable due to the presence of C-forests forming a hedge.

We now prove that **ID** terminates and is sound.

Lemma 2 ***ID** exhausts all valid input, and always terminates.*

Proof: First we show that **ID** is exhaustive for any valid input. All recursive calls to **ID** pass a non-empty graph as arguments, so third and fourth lines are exhaustive. Given this, we must show that any input reaching the last two lines will satisfy one of the two corresponding conditions. Assume $C(G \setminus \mathbf{X}) = \{S\}$. Then if there are no bidirected arcs in G from S to $\mathbf{X}$, the penultimate condition holds. Otherwise, the ultimate condition holds.

Now the result follows if any recursive call of **ID** reduces either the size of the set $\mathbf{X}$ or the size of the set $\mathbf{V} \setminus \mathbf{X}$. This is only non-obvious for the last recursive call. However, if the last recursive call is made, then $(\exists X \in \mathbf{X}) X \not\in S'$, or the failure condition would have been triggered. □

To show soundness, we need a number of utility lemmas justifying various lines of the algorithm. In the proofs, we will utilize the 3 rules of do-calculus. These rules allow insertion and deletion of interventions and observational evidence into and from distributions, using probabilistic independencies implied by the causal graph. Do-calculus is discussed in [Pearl, 2000]. The 3 rules themselves are reproduced in the last section of the paper.

First, we must show that an effect of the form $P_{\mathbf{x}}(\mathbf{y})$ decomposes according to the set of C-components of the graph $G \setminus \mathbf{X}$.

Lemma 3 *Let M be a causal model with graph G. Let $\boldsymbol{y}, \boldsymbol{x}$ be value assignments. Let $C(G \setminus \boldsymbol{X}) = \{S_1, ..., S_k\}$. Then $P_{\boldsymbol{x}}(\boldsymbol{y}) = \sum_{\boldsymbol{v} \setminus (\boldsymbol{y} \cup \boldsymbol{x})} \prod_i P_{\boldsymbol{v} \setminus s_i}(s_i)$.*

Proof: There is a simple proof for this based on C-component factorizations of P. Fix a model M with graph G and assume $C(G) = \{S'_1, ..., S'_k\}$. Then $P(\mathbf{v}) = \prod_i P_{\mathbf{v} \setminus s'_i}(s'_i)$. The same factorization applies to the submodel $M_{\mathbf{x}}$ which induces the graph $G \setminus \mathbf{X}$, which implies the result. However, for the purposes of the next section, we also provide a longer proof using do-calculus, which can be found in the Appendix. □

Next we show that to identify the effect on $\mathbf{Y}$, it is sufficient to restrict our attention to the ancestor set of $\mathbf{Y}$.

Lemma 4 *Let $\boldsymbol{Y}, \boldsymbol{X}$ be sets of variables. Let M be a causal model with graph G. Then $P_{\boldsymbol{x}}(\boldsymbol{Y})$ is identifiable in G from P if and only if $P_{\boldsymbol{x}}(\boldsymbol{Y})$ is identifiable in $G_{An(\boldsymbol{Y})}$ from $\sum_{\boldsymbol{v} \setminus An(\boldsymbol{Y})_G} P$.*

Proof: This follows from factorization properties of $P(\mathbf{V})$. See [Tian, 2002]. □

Next we prove soundness for line 3.

Lemma 5 *Let $\boldsymbol{W} = (\boldsymbol{V} \setminus \boldsymbol{X}) \setminus An(\boldsymbol{Y})_{G_{\overline{\boldsymbol{x}}}}$. Then $P_{\boldsymbol{x}}(\boldsymbol{y}) = P_{\boldsymbol{x},\boldsymbol{w}}(\boldsymbol{y})$, where $\boldsymbol{w}$ are arbitrary values of $\boldsymbol{W}$.*

Proof: Note that by assumption, $\mathbf{Y} \perp\!\!\!\perp \mathbf{W} | \mathbf{X}$ in $G_{\overline{\mathbf{x}},\overline{\mathbf{w}}}$. The conclusion follows by rule 3 of do-calculus (see next section). □

Finally, we show the soundness of the last recursive call.

Lemma 6 *Let $\boldsymbol{Y}, \boldsymbol{X}$ be sets of variables. Let M be a causal model with graph G. Whenever the conditions of the last recursive call are satisfied, $P_{\boldsymbol{x}}$ is computable in G from P, if and only if $P_{\boldsymbol{x} \cap S'}$ is computable in S' from $P' = \prod_{\{i | V_i \in S'\}} P(V_i | V_G^{(i-1)} \cap S', v_G^{(i-1)} \setminus S')$.*

Proof:

Note that when the last recursive call executes, $\mathbf{X}$ and S partition G. We must show that $P_{\mathbf{x}}(\mathbf{y})$ is identifiable from $P(\mathbf{V})$ in G if and only if $P_{\mathbf{x} \cap S'}(\mathbf{y})$ is identifiable from $P_{\mathbf{x}}(S')$ in S'. Then it is not difficult to show $P_{\mathbf{x}}(S') = Q[S'] = \prod_{\{i | V_i \in S'\}} P(V_i | V_G^{(i-1)} \cap S', v_G^{(i-1)} \setminus S')$.

Note that for any causal model M with graph G, whenever $\mathbf{X}$ and S partition G, the submodel $M_{\mathbf{x} \setminus S'}(S')$ induces precisely the graph S'. If $P_{\mathbf{x}}(\mathbf{y})$ is identifiable from $P_{\mathbf{x} \cap S'}(\mathbf{y})$, then it is certainly identifiable from $P(\mathbf{V})$, since as we saw from the previous Lemma, $P_{\mathbf{x} \cap S'}(\mathbf{y})$ can be expressed in terms of from $P(\mathbf{V})$. If $P_{\mathbf{x}}(\mathbf{y})$ is not identifiable from $P_{\mathbf{x} \cap S'}(\mathbf{y})$ in S', then certainly it is not identifiable in G from $P(\mathbf{V})$, because adding to the graph cannot help identifiability. Finally, because $M_{\mathbf{x} \setminus S'}(S')$ already fixes $\mathbf{x} \setminus S'$, $P_{\mathbf{x}}(\mathbf{y})$ in M is equal to $P_{\mathbf{x} \cap S'}(\mathbf{y})$ in $M_{\mathbf{x} \setminus S'}$. □

We can now show the soundness of **ID**.

Theorem 5 (soundness) *Whenever **ID** returns an expression for $P_{\boldsymbol{x}}(\boldsymbol{y})$, it is correct.*

Proof: The soundness of all recursive calls has already been shown. What remains is to show the soundness of all positive base cases. If $\mathbf{x} = \emptyset$, the desired effect can be obtained from P by marginalization, thus this base case is clearly correct. If $C(G \setminus \mathbf{X}) = \{S\}$ and $S \in C(G)$, correctness follows from the properties of C-components, see [Tian, 2002]. □

We can now characterize the relationship between C-forests, the set $\mathbf{Y}$ and the failure of **ID** to return an expression.

Theorem 6 *Assume **ID** fails to identify $P_{\boldsymbol{x}}(\boldsymbol{y})$ (executes line 5). Then there exist $\boldsymbol{X}' \subseteq \boldsymbol{X}$, $\boldsymbol{Y}' \subseteq \boldsymbol{Y}$ such that the graph pair G, S returned by the fail condition of **ID** contain as edge subgraphs C-forests F, F' that form a hedge for $P_{\boldsymbol{x}'}(\boldsymbol{y}')$.*

Proof: Consider line 5, and G and $\mathbf{y}$ local to that recursive call. Let $\mathbf{R}$ be the root set of G. Since G is a single C-component, it is possible to remove a set of directed arrows from G while preserving the root set $\mathbf{R}$ such that the resulting graph F is an $\mathbf{R}$-rooted C-forest.

Moreover, since $F' = F \cap S$ is closed under descendants, and since only single directed arrows were removed from S to obtain F', F' is also a C-forest. $F' \cap \mathbf{X} = \emptyset$, and $F \cap \mathbf{X} \neq \emptyset$ by construction. $\mathbf{R} \subseteq An(\mathbf{Y})_{G_{\overline{\mathbf{x}}}}$ by lines 2 and 3 of the algorithm. It's also clear that $\mathbf{y}, \mathbf{x}$ local to the recursive call in question are subsets of the original input. □

Corollary 2 (completeness) ***ID** is complete.*

Proof: By the previous Theorem, if **ID** fails, then $P_{\mathbf{x}'}(\mathbf{y}')$ is not identifiable in a subgraph H of G. Moreover, $\mathbf{X} \cap H = \mathbf{X}'$, by construction of H. As such, it is easy to extend the counterexamples in the previous Theorem with variables independent of H, with the resulting models inducing G, and witnessing the non-identifiability of $P_{\mathbf{x}}(\mathbf{y})$. □

The following is now immediate.

Corollary 3 *$P_{\boldsymbol{x}}(\boldsymbol{y})$ is identifiable from P in G if and only if there does not exist a hedge for $P_{\boldsymbol{x}'}(\boldsymbol{y}')$ in G, for any $\boldsymbol{X}' \subseteq \boldsymbol{X}$ and $\boldsymbol{Y}' \subseteq \boldsymbol{Y}$.*

So far we have not only establishes completeness, but also fully characterized graphically all situations where joint interventional distributions are identifiable. We can use these results to derive a characterization of identifiable models, that is, causal models where all interventional distributions are identifiable.

Corollary 4 (model identification) *Let G be a semi-Markovian causal diagram. Then all interventional distributions are identifiable in G if and only if G does not contain an $\boldsymbol{R}$-rooted C-forest F such that $F \setminus \boldsymbol{R} \neq \emptyset$.*

Proof: Note that if F, F' are C-forests which form a hedge for some effect, there must be a variable $X \in F$, which is an ancestor of another variable $Y \in F$. Therefore if G does not contain any $\mathbf{R}$-rooted C-forests F such that $F \setminus \mathbf{R} \neq \emptyset$, then **ID** never reaches the fail condition. Thus all effects are identifiable.

Otherwise, let F be such a forest. Fix $X, Y \in F$ such that $X \in Pa(Y)_F$. Note that X and Y are connected by a path consisting entirely of bidirected arcs, by definition of C-forests. Let $W_1, ..., W_k \in F$ be the set of nodes on this path. It is not hard to see that $P_x(Y, W_1, ..., W_k)$ is not identifiable in G. □

Connections to Existing Identification Algorithms

In the previous section we established that **ID** is a sound and complete algorithm for all unconditional effects of the form

$P_{\mathbf{x}}(\mathbf{y})$. It is natural to ask whether this result can be used to show completeness of earlier algorithms conjectured to be complete.

First, we consider a declarative algorithm known as do-calculus [Pearl, 2000], which has been a popular technique for proving identifiability, with its completeness remaining an open question. Do-calculus allows one to manipulate expressions corresponding to effects using the following three identities:

- Rule 1: $P_{\mathbf{x}}(\mathbf{y}|\mathbf{z},\mathbf{w}) = P_{\mathbf{x}}(\mathbf{y}|\mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z}|\mathbf{X},\mathbf{W})_{G_{\overline{\mathbf{X}}}}$
- Rule 2: $P_{\mathbf{x},\mathbf{z}}(\mathbf{y}|\mathbf{w}) = P_{\mathbf{x}}(\mathbf{y}|\mathbf{z},\mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z}|\mathbf{X},\mathbf{W})_{G_{\overline{\mathbf{X}},\underline{\mathbf{Z}}}}$
- Rule 3: $P_{\mathbf{x},\mathbf{z}}(\mathbf{y}|\mathbf{w}) = P_{\mathbf{x}}(\mathbf{y}|\mathbf{w})$ if $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z}|\mathbf{X},\mathbf{W})_{G_{\overline{X},\overline{Z(\mathbf{W})}}}$

where $Z(\mathbf{W}) = \mathbf{Z} \setminus An(\mathbf{W})_{G_{\overline{\mathbf{X}}}}$.

We show that the steps of the algorithm **ID** correspond to sequences of standard probabilistic manipulations, and applications of rules of do-calculus, which entails completeness of do-calculus for identifying unconditional effects.

Theorem 7 *The rules of do-calculus, together with standard probability manipulations are complete for determining identifiability of all effects of the form $P_{\mathbf{x}}(\mathbf{y})$.*

Proof outline: We must show that all operations correspondings to lines of **ID** correspond to sequences of standard probability manipulations and applications of the rules of do-calculus. These manipulations are done either on the effect expression $P_{\mathbf{x}}(\mathbf{y})$, or the observational distribution P, until the algorithm either fails, or the two expressions 'meet' by producing a single chain of manipulations.

Line 1 is just standard probability operations.

Line 2 follows by rule 1 of do-calculus, coupled with the semi-Markovian property.

Line 3 corresponds to an application of rule 3 of do-calculus by Lemma 5.

Line 4 corresponds to a sequence of do-calculus rule applications and probability manipulations by Lemma 3.

Line 5 is a fail condition.

Line 6 implies that G is partitioned into sets S and $\mathbf{X}$, and there no bidirected arcs from $\mathbf{X}$ to S. This implies that $S \perp\!\!\!\perp \mathbf{X}|\mathbf{X}$ in $G_{\underline{\mathbf{X}}}$. The operation of line 6 then corresponds to an application of Rule 2 of do-calculus.

To prove line 7, we note that since all nodes not in S' are in $\mathbf{X}$ at the time of the last recursive call, and since $\mathbf{X} \setminus S'$ does not contain any bidirected arcs to S', we can conclude that $\prod_{\{i|V_i \in S'\}} P(V_i|V_G^{(i-1)} \cap S', v_G^{(i-1)} \setminus S') = P_{\mathbf{x} \setminus S'}(S')$ by Rule 2, by the same reasoning used to prove the result for line 6. □

Next, we consider a version of an identification algorithm due to Tian, shown in Fig. 5. The soundness of this algorithm has already been addressed elsewhere, so we turn to the matter of completeness.

Theorem 8 *Assume **jointident** fails to identify $P_{\mathbf{x}}(\mathbf{y})$. Then there exist C-forests F, F' forming a hedge for $P_{\mathbf{x}'}(\mathbf{y}')$, where $\mathbf{X}' \subseteq \mathbf{X}$, $\mathbf{Y}' \subseteq \mathbf{Y}$.*

Proof: Assume **identify** fails. Consider C-components C, T local to the failed recursive call. Let $\mathbf{R}$ be the root set of C. Because $T = An(C)_{G_T}$, $\mathbf{R}$ is also a root set of T. As in the proof

function **identify**(C, T, Q[T])
INPUT: $C \subseteq T$, both are C-components, $Q[T]$ a probability distribution
OUTPUT: Expression for $Q[C]$ in terms of $Q[T]$ or **FAIL**

Let $A = An(C)_{G_T}$.

1 If $A = C$, return $\sum_{T \setminus C} P$

2 If $A = T$, return **FAIL**

3 If $C \subset A \subset T$, there exists a C-component T' such that $C \subset T' \subset A$.
return **identify**$(C, T', Q[T'])$
($Q[T']$ is known to be computable from $\sum_{T \setminus A} Q[T]$)

function **jointident**(**y**, **x**, P, G)
INPUT: **x**,**y** value assignments, P a probability distribution, G a causal diagram.
OUTPUT: Expression for $P_{\mathbf{x}}(\mathbf{y})$ in terms of P or **FAIL**.

1 Let $D = An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$.

2 Assume $C(D) = \{D_1, ..., D_k\}$, $C(G) = \{C_1, ..., C_m\}$.

3 return $\sum_{D \setminus S} \prod_i$ **identify**$(D_i, C_{D_i}, Q[C_{D_i}])$,
where $(\forall i) D_i \subseteq C_{D_i}$

Figure 5: An identification algorithm modified from [Tian, 2002]

of Theorem 6, we can remove a set of directed arrows from C and T while preserving $\mathbf{R}$ as the root set such that the resulting edge subgraphs are C-forests. By line 1 of **jointident**, $C, T \subseteq An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$.

Finally, because **identify** will always succeed if $D_i = C_{D_i}$, it must be the case that $D_i \subset C_{D_i}$. But this implies $\mathbf{X} \cap C = \emptyset$, $\mathbf{X} \cap T \neq \emptyset$. Thus, edge subsets of C, and T satisfy the properties of a hedge for $P_{\mathbf{x}'}(\mathbf{y}')$, where $\mathbf{X}' \subseteq \mathbf{X}$, $\mathbf{Y}' \subseteq \mathbf{Y}$. □

Corollary 5 ***jointident** is complete.*

Proof: This is implied by the previous Theorem, and Corollary 3. □

Appendix

In this appendix we present the full proofs of Theorems 3 and 4, and Lemma 3

Theorem 3 *Let G be a Y-rooted C-tree. Then the effects of any set of nodes in G on Y are not identifiable.*

The proof will proceed by constructing a family of counterexamples. For any such G and any set of nodes $\mathbf{X}$, we will construct two causal models M^1 and M^2 that will agree on the joint distribution over endogenous variables, but disagree on the interventional distribution $P_{\mathbf{x}}(Y)$. The argument will be a strict generalization of the proof of Theorem 1.

The two models in question agree on the following features. All variables in $\mathbf{U} \cup \mathbf{V}$ are binary. All exogenous variables are distributed uniformly. All endogeous variables except Y are set to the bit parity of the values of their parents.

The only difference between the models is how the function for Y is defined. Let $bp(.)$ denote the bit parity (sum mod 2) function. Let the function $f_Y : U, Pa(Y) \to Y$ be defined as follows.

M^1: $f_Y(u, pa(Y)) = bp(u, pa(Y))$.

M^2: $f_Y(u, pa(Y)) = 0$.

We will prove the necessary claims about M^1 and M^2 in the following lemmas.

Lemma 7 *The observational distributions in the two models are the same. In other words* $P^1(\mathbf{V}) = P^2(\mathbf{V})$.

Proof: Since the two models agree on $P(\mathbf{U})$ and all functions except f_Y, it suffices to show f_Y maintains the same input/output behavior in both models. Clearly, Y always outputs 0 in M^2.

Consider the behavior of f_Y in M^1. We must show that the bit parity of the parents of Y is always 0. Since the endogenous nodes and directed edges form a tree with Y at the root, and since the functions for all nodes compute the bit parity, or sum $(\bmod 2)$, of their parents we can view Y as computing the sum $(\bmod 2)$ of all values of nodes in $\mathbf{U}$. Moreover, since each variable in $\mathbf{U}$ has exactly two endogenous children, the bit value of any node in $\mathbf{U}$ is counted twice. Thus, regardless of the value assignment of all nodes $\mathbf{U}$, the sum $(\bmod 2)$ computed at Y will be 0. □

Lemma 8 *For any set of variables* $\mathbf{X}$ *(excluding* Y*), there exists a value assignment* $\mathbf{x}$ *such that* $P^1_{\mathbf{x}}(Y) \neq P^2_{\mathbf{x}}(Y)$.

Proof: Consider any such set $\mathbf{X}$. Let $\mathbf{A} = An(\mathbf{X})$, $\mathbf{B} = G \setminus \mathbf{A}$. Note that $\mathbf{X} \subseteq \mathbf{A}$, $Y \in \mathbf{B}$, so both sets are non-empty. By definition of C-trees, there exists a bidirected arc between a node in $\mathbf{A}$ and a node in $\mathbf{B}$. Because the corresponding variable U will only have its value counted once for the purposes of bit parity, $P^1_{\mathbf{x}}(Y = 1) > 0$. By construction, $P^2_{\mathbf{x}}(Y = 1) = 0$. □

Lemmas 7 and 8 together with the fact that in our counterexamples any two disjoint subsets of $\mathbf{U}$ are independent, prove Theorem 3. It is straightforward to generalize this proof for the positive $P(\mathbf{V})$ in the same way as Theorem 1.

Theorem 4 *Assume* F, F' *form a hedge for* $P_{\mathbf{x}}(\mathbf{y})$*. Then* $P_{\mathbf{x}}(\mathbf{y})$ *is not identifiable.*

We will consider counterexamples with the induced graph $H = De(F)_G \cap An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$. Without loss of generality, assume H is a forest. Once again, we generalize the proof for the bow-arc graphs and C-trees. As before, we have two models with binary nodes. In the first model, the values of all observable nodes are equal to the bit parity of the parent values. In the second model, the same is true, except any node in F' disregards the parent values if the parent is in F.

Lemma 9 *Given the models defined above,* $P^1(\mathbf{V}) = P^2(\mathbf{V})$.

Proof: Let F be an $\mathbf{R}$-rooted C-forest, let $\mathbf{V}'$ be the set of observable variables and $\mathbf{U}'$ be the set of unobservable variables in F. Assume the functions of $\mathbf{V}'$ and distributions of $\mathbf{U}'$ are defined as above. As a subclaim, we want to show that any node assignment where the bit parity of $\mathbf{R}$ is odd has probability 0, and all other assignments are equally likely.

As in the proof of Lemma 7, we can view $\mathbf{R}$ as computing the bit parity of variables $\mathbf{U}'$. Moreover, because variable in $\mathbf{U}'$ has exactly two endogenous children in F, each such variable is counted twice. Thus, the bit parity of $\mathbf{R}$ is always even.

Next, we want to prove that the function from $\mathbf{U}'$ to $\mathbf{V}'$ is 1-1. Assume this is not so, and fix two instantiations of $\mathbf{U}'$ that map to the same values in $\mathbf{V}'$, and that differ by the settings of the set $\mathbf{U}^* = \{U_1, ..., U_k\}$. Since bidirected edges form a spanning tree there exist some nodes in $\mathbf{V}^*$ with an odd number of parents in $\mathbf{U}^*$. Order all such nodes topologically in G, and consider the topmost node, call it X. Clearly, if we flip all values in $\mathbf{U}^*$, and no other values in $\mathbf{U}'$, the value of X will also flip. Contradiction.

Since all $\mathbf{U}'$ are uniformly distributed, and $|\mathbf{U}'| + 1 = |\mathbf{V}'|$, our subclaim follows.

Let $\mathbf{D} = F \setminus F'$. All we have left to show is that in M^2 all value assignments over $\mathbf{D}$ are equally likely. Consider the distribution $P^1(\mathbf{D})$. Because P^1 decomposes as: $\prod_{V \in G} P(V|Pa(V))$, where $Pa(V)$ includes variables in $\mathbf{U}'$, it's easy to see that $P^1(\mathbf{D})$ decomposes as $\prod_{V \in G_{\mathbf{D}}} P(V|Pa(V))$. But this implies $P^1(\mathbf{D}) = P^2(\mathbf{D})$.

Since we already know $P^1(F') = P^2(F')$, and because the two models agree on all functions in $H \setminus F$, we obtain our result. □

Lemma 10 *For any* $\mathbf{x}$, $P(\mathbf{x}|Pa(\mathbf{X})_G \setminus \mathbf{X}) > 0$.

Proof: In the last lemma we established that all value assignments where the bit parity of $\mathbf{y}$ is even are equally likely, and all value assignments where the bit parity of $\mathbf{y}$ is odd are impossible. This implies that $\sum_{\mathbf{y}} P(\mathbf{V})$ is a positive distribution. Our result follows. □

Lemma 11 *There exists a value assignment* $\mathbf{x}$ *to variables* $\mathbf{X}$ *such that* $P^1_{\mathbf{x}}(\mathbf{R}) \neq P^2_{\mathbf{x}}(\mathbf{R})$.

Proof: This is an easy generalization of Lemma 8. As before, we can find a variable U with a value counted once for the purposes of bit parity in $M^1_{\mathbf{x}}$. This implies that $P^1_{\mathbf{x}}(\sum \mathbf{y} \pmod 2 = 1) > 0$, while $P^2_{\mathbf{x}}(\sum \mathbf{y} \pmod 2 = 1) = 0$ by construction. □

Lemma 12 *If* $P_{\mathbf{x}}(\mathbf{R})$ *is not identifiable, then neither is* $P_{\mathbf{x}}(\mathbf{Y})$.

Proof: We observe that the root set in H is a subset of $\mathbf{Y}$, call it $\mathbf{Y}'$. Since $P^1(\mathbf{V}') = P^2(\mathbf{V}')$, and both models agree on the functions in $H \setminus F$, we obtain that $P^1(\mathbf{V}) = P^2(\mathbf{V})$ in H. Since H is a forest, it is not difficult to show that the bit parity of $\mathbf{R}$ is equal to the bit parity of $\mathbf{Y}'$ in both submodels $M^1_{\mathbf{x}}, M^2_{\mathbf{x}}$ for any $\mathbf{x}$. This completes the proof. □

The previous lemmas together prove Theorem 4. We now provide a proof of Lemma 3 using the rules of do-calculus.

Lemma 3 *Let* M *be a causal model with graph* G*. Let* $\mathbf{y}, \mathbf{x}$ *be value assignments. Let* $C(G \setminus \mathbf{X}) = \{S_1, ..., S_k\}$*. Then* $P_{\mathbf{x}}(\mathbf{y}) = \sum_{\mathbf{v} \setminus (\mathbf{y} \cup \mathbf{x})} \prod_i P_{\mathbf{v} \setminus s_i}(s_i)$.

Proof: The identities

$$P_{\mathbf{x}}(\mathbf{y}) = \sum_{\mathbf{v}\setminus(\mathbf{x}\cup\mathbf{y})} P_{\mathbf{x}}(\mathbf{v}) = \sum_{\mathbf{v}\setminus(\mathbf{x}\cup\mathbf{y})} \prod_i P_{\mathbf{x}}(v_i|v_G^{(i-1)}) =$$

$$\sum_{\mathbf{v}\setminus(\mathbf{x}\cup\mathbf{y})} \prod_{\{j|S_j\in C(G)\}} \prod_{\{i|V_i\in S_j\}} P_{\mathbf{x}}(v_i|v_G^{(i-1)}) =$$

are standard probability manipulations licensed by the structure of the graph.

$$\sum_{\mathbf{v}\setminus(\mathbf{x}\cup\mathbf{y})} \prod_{\{j|S_j\in C(G)\}} \prod_{\{i|V_i\in S_j\}} P_{\mathbf{x},v_G^{(i-1)}\setminus S_j}(v_i|v_G^{(i-1)} \cap S_j) =$$

For this identity, fix V_i, and consider a vertex $A \in V_G^{(i-1)} \setminus S_j$, where $V_i \in S_j$. A backdoor path from A to V_i must involve only ancestors of V_i since only such nodes are observed in our expression, and since our graph is acyclic. A backdoor path also cannot involve singly directed edges, since all ancestors of V_i are observed, blocking all such paths. A backdoor path also cannot involve only bidirected arcs, since $A \not\in S_j$. Thus, no such path exists, and the identity follows from Rule 2 of do-calculus. Note that any ancestor of S_j not itself in S_j is now fixed by intervention.

$$\sum_{\mathbf{v}\setminus(\mathbf{x}\cup\mathbf{y})} \prod_{\{j|S_j\in C(G)\}} P_{\mathbf{x},\mathbf{v}\setminus s_j}(s_j) = \sum_{\mathbf{v}\setminus(\mathbf{x}\cup\mathbf{y})} \prod_{\{j|S_j\in C(G\setminus\mathbf{X})\}} P_{\mathbf{v}\setminus s_j}(s_j)$$

.

The last two identities add interventions for non-ancestors of S_j (using Rule 3 of do-calculus), and group terms using rules of probability. □

Conclusions

We have presented a graphical structure called a hedge which we used to completely characterize situations when joint interventional distributions are identifiable in semi-Markovian causal models. We were then able to use this characterization to construct a sound and complete identification algorithm, and prove completeness of two existing algorithms for the same identification problem.

The natural open question stemming from this work is whether the improved algorithm presented can lead to the identification of $conditional\ interventional\ distributions$ of the form $P_{\mathbf{x}}(\mathbf{y}|\mathbf{z})$. In addition, complete conditions for identification of experiments may open the door to the identification of more complex counterfactual quantities in semi-Markovian models, like path-specific effects [Avin, Shpitser, & Pearl, 2005], or even first-order counterfactual statements.

References

[Avin, Shpitser, & Pearl, 2005] Avin, C.; Shpitser, I.; and Pearl, J. 2005. Identifiability of path-specific effects. In *International Joint Conference on Artificial Intelligence*, volume 19, 357–363.

[Halpern, 2000] Halpern, J. 2000. Axiomatizing causal reasoning. *Journal of A.I. Research* 317–337.

[Huang & Valtorta, 2006] Huang, Y., and Valtorta, M. 2006. On the completeness of an identifiability algorithm for semi-markovian models. Technical Report TR-2006-01, Computer Science and Engineering Department, University of South Carolina.

[Kuroki & Miyakawa, 1999] Kuroki, M., and Miyakawa, M. 1999. Identifiability criteria for causal effects of joint interventions. *Journal of Japan Statistical Society* 29:105–117.

[Pearl & Robins, 1995] Pearl, J., and Robins, J. M. 1995. Probabilistic evaluation of sequential plans from causal models with hidden variables. In *Uncertainty in Artificial Intelligence*, volume 11, 444–453.

[Pearl, 1995] Pearl, J. 1995. Causal diagrams for empirical research. *Biometrika* 82(4):669–709.

[Pearl, 2000] Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

[Pearl, 2001] Pearl, J. 2001. Direct and indirect effects. In *Proceedings of UAI-01*, 411–420.

[Spirtes, Glymour, & Scheines, 1993] Spirtes, P.; Glymour, C.; and Scheines, R. 1993. *Causation, Prediction, and Search*. Springer Verlag, New York.

[Tian & Pearl, 2002] Tian, J., and Pearl, J. 2002. A general identification condition for causal effects. In *Eighteenth National Conference on Artificial Intelligence*, 567–573.

[Tian, 2002] Tian, J. 2002. *Studies in Causal Reasoning and Learning*. Ph.D. Dissertation, Department of Computer Science, University of California, Los Angeles.