In Proceedings of the Eighteenth National Conference on Artificial Intelligence, AAAI Press/The MIT Press: Menlo Park, CA, 567-573, August 2002.

TECHNICAL REPORT
R-290-A
August 2002

A General Identification Condition for Causal Effects

Jin Tian and Judea Pearl
Cognitive Systems Laboratory
Computer Science Department
University of California, Los Angeles, CA 90024
{jtian, judea}@cs.ucla.edu

Abstract

This paper concerns the assessment of the effects of actions or
policy interventions from a combination of: (i) nonexperimental data, and (ii) substantive assumptions. The assumptions are encoded in the form of a directed acyclic graph, also called "causal graph", in which some variables are presumed to be unobserved. The paper establishes a necessary and sufficient criterion for the identifiability of the causal effects of a singleton variable on all other variables in the model, and a powerful sufficient criterion for the effects of a singleton variable on any set of variables.

Introduction

This paper explores the feasibility of inferring cause effect relationships from various combinations of data and theoretical assumptions. The assumptions considered will be represented in the form of an acyclic causal diagram which contains both arrows and bi-directed arcs (Pearl 1995; 2000). The arrows represent the potential existence of direct causal relationships between the corresponding variables, and the bi-directed arcs represent spurious dependencies due to unmeasured confounders. Our main task will be to decide whether the assumptions represented in any given diagram are sufficient for assessing the strength of causal effects from nonexperimental data and, if sufficiency is proven, to express the target causal effect in terms of estimable quantities.

It is well known that, in the absence of unmeasured confounders, all causal effects are *identifiable*, that is, the joint response of any set Y of variables to intervention on a set T of treatment variables, denoted $P_t(y)$,[1] can be estimated consistently from nonexperimental data (Robins 1987; Spirtes, Glymour, & Scheines 1993; Pearl 1993). If some confounders are not measured, then the question of identifiability arises, and whether the desired quantity can be estimated depends critically on the precise locations (in the diagram) of those confounders vis a vis the sets T and Y. Sufficient graphical conditions for ensuring the identification of $P_t(y)$ were established by several authors (Spirtes, Glymour, & Scheines 1993; Pearl 1993; 1995) and are summarized in (Pearl 2000, Chapters 3 and 4). For example, a criterion called "back-door" permits one to determine whether a given causal effect $P_t(y)$ can be obtained by "adjustment", that is, whether a set C of covariates exists such that

$$P_t(y) = \sum_c P(y|c,t)P(c) \tag{1}$$

When there exists no set of covariates that is sufficient for adjustment, causal effects can sometimes be estimated by invoking multi-stage adjustments, through a criterion called "front-door" (Pearl 1995). More generally, identifiability can be decided using do-calculus derivations (Pearl 1995), that is, a sequence of syntactic transformations capable of reducing expressions of the type $P_t(y)$ to subscript-free expressions. Using do-calculus as a guide, (Galles & Pearl 1995) devised a graphical criterion for identifying $P_x(y)$ (where X and Y are singletons) that combines and expands the "front-door" and "back-door" criteria (see (Pearl 2000, pp. 114-8)).

This paper develops new graphical identification criteria that generalize and simplify existing criteria in several ways. We show that $P_x(v)$, where X is a singleton and V is the set of all variables excluding X, is identifiable if and only if there is no consecutive sequence of confounding arcs between X and X's immediate successors in the diagram.[2] When interest lies in the effect of X on a subset S of outcome variables, not on the entire set V, it is possible that $P_x(s)$ would be identifiable even though $P_x(v)$ is not. To this end, the paper gives a sufficient criterion for identifying $P_x(s)$, which is an extension of the criterion for identifying $P_x(v)$. It says that $P_x(s)$ is identifiable if there is no consecutive sequence of confounding arcs between X and X's children in the subgraph composed of the ancestors of S. Other than this requirement, the diagram may have an arbitrary structure, including any number of confounding arcs between X and S. This simple criterion is shown to cover all criteria reported in the literature (with X singleton), including the "back-door", "front-door", and those developed by (Galles & Pearl 1995).

[1](Pearl 1995; 2000) used the notation $P(y|set(t))$, $P(y|do(t))$, or $P(y|\hat{t})$ for the post-intervention distribution, while (Lauritzen 2000) used $P(y||t)$.

[2]A variable Z is an "immediate successor" (or a "child") of X if there exists an arrow $X \to Z$ in the diagram.

Notation, Definitions, and Problem Formulation

The use of causal models for encoding distributional and causal assumptions is now fairly standard (see, for example, (Pearl 1988; Spirtes, Glymour, & Scheines 1993; Greenland, Pearl, & Robins 1999; Lauritzen 2000; Pearl 2000)). The simplest such model, called *Markovian*, consists of a directed acyclic graph (DAG) over a set $V = \{V_1, \ldots, V_n\}$ of vertices, representing variables of interest, and a set E of directed edges, or arrows, that connect these vertices. The interpretation of such a graph has two components, probabilistic and causal. The probabilistic interpretation views the arrows as representing probabilistic dependencies among the corresponding variables, and the missing arrows as representing conditional independence assertions: Each variable is independent of all its nondescendants given its direct parents in the graph.[3] These assumptions amount to asserting that the joint probability function $P(v) = P(v_1, \ldots, v_n)$ factorizes according to the product

$$P(v) = \prod_i P(v_i|pa_i) \tag{2}$$

where pa_i are (values of) the parents of variable V_i in the graph.[4]

The causal interpretation views the arrows as representing causal influences between the corresponding variables. In this interpretation, the factorization of (2) still holds, but the factors are further assumed to represent autonomous data-generation processes, that is, each conditional probability $P(v_i|pa_i)$ represents a stochastic process by which the values of V_i are chosen in response to the values pa_i (previously chosen for V_i's parents), and the stochastic variation of this assignment is assumed independent of the variations in all other assignments. Moreover, each assignment process remains invariant to possible changes in the assignment processes that govern other variables in the system. This modularity assumption enables us to predict the effects of interventions, whenever interventions are described as specific modifications of some factors in the product of (2). The simplest such intervention involves fixing a set T of variables to some constants $T = t$, which yields the post-intervention distribution

$$P_t(v) = \begin{cases} \prod_{\{i|V_i \notin T\}} P(v_i|pa_i) & v \text{ consistent with } t. \\ 0 & v \text{ inconsistent with } t. \end{cases} \tag{3}$$

Eq. (3) represents a truncated factorization of (2), with factors corresponding to the manipulated variables removed. This truncation follows immediately from (2) since, assuming modularity, the post-intervention probabilities $P(v_i|pa_i)$ corresponding to variables in T are either 1 or 0, while those corresponding to unmanipulated variables remain unaltered.[5] If T stands for a set of treatment variables and Y for an outcome variable in $V \setminus T$, then Eq. (3) permits us to calculate the probability $P_t(y)$ that event $Y = y$ would occur if treatment condition $T = t$ were enforced uniformly over the population. This quantity, often called the *causal effect* of T on Y, is what we normally assess in a controlled experiment with T randomized, in which the distribution of Y is estimated for each level t of T.

We see from Eq. (3) that the model needed for predicting the effect of interventions requires the specification of three elements

$$M = \langle V, G, P(v_i|pa_i) \rangle$$

where (i) $V = \{V_1, \ldots, V_n\}$ is a set of variables, (ii) G is a directed acyclic graph with nodes corresponding to the elements of V, and (iii) $P(v_i|pa_i), i = 1, \ldots, n$, is the conditional probability of variable V_i given its parents in G. Since $P(v_i|pa_i)$ is estimable from nonexperimental data whenever V is observed, we see that, given the causal graph G, all causal effects are estimable from the data as well.[6]

Our ability to estimate $P_t(v)$ from nonexperimental data is severely curtailed when some variables in a Markovian model are unobserved, or, equivalently, if two or more variables in V are affected by unobserved confounders; the presence of such confounders would not permit the decomposition in (2). Let V and U stand for the sets of observed and unobserved variables, respectively. Assuming that no U variable is a descendant of any V variable (called a *semi-Markovian* model), the observed probability distribution, $P(v)$, becomes a mixture of products:

$$P(v) = \sum_u \prod_i P(v_i|pa_i, u^i) P(u) \tag{4}$$

where pa_i and u^i stand for the sets of the observed and unobserved parents of V_i, and the summation ranges over all the U variables. The post-intervention distribution, likewise, will be given as a mixture of truncated products

$$P_t(v) = \begin{cases} \sum_u \prod_{\{i|V_i \notin T\}} P(v_i|pa_i, u^i) P(u) & v \text{ consistent with } t. \\ 0 & v \text{ inconsistent with } t. \end{cases} \tag{5}$$

and, the question of identifiability arises, i.e., whether it is possible to express $P_t(v)$ as a function of the observed distribution $P(v)$.

Formally, our semi-Markovian model consists of five elements

$$M = \langle V, U, G_{VU}, P(v_i|pa_i, u^i), P(u) \rangle$$

[3]We use family relationships such as "parents," "children," "ancestors," and "descendants," to describe the obvious graphical relationships. For example, the parents PA_i of node V_i are the set of nodes that are directly connected to V_i via arrows pointing to V_i.

[4]We use uppercase letters to represent variables or sets of variables, and use corresponding lowercase letters to represent their values (instantiations).

[5]Eq. (3) was named "Manipulation Theorem" in (Spirtes, Glymour, & Scheines 1993), and is also implicit in Robins' (1987) G-computation formula.

[6]It is in fact enough that the parents of each variable in T be observed (Pearl 2000, p. 78).

where G_{VU} is a causal graph consisting of variables in $V \times U$. Clearly, given M and any two sets T and S in V, $P_t(s)$ can be determined unambiguously using (5). The question of identifiability is whether a given causal effect $P_t(s)$ can be determined uniquely from the distribution $P(v)$ of the observed variables, and is thus independent of the unknown quantities, $P(u)$ and $P(v_i|pa_i, u^i)$, that involve elements of U.

In order to analyze questions of identifiability, it is convenient to represent our modeling assumptions in the form of a graph G that does not show the elements of U explicitly but, instead, represents the confounding effects of U using bidirected edges. A bidirected edge between nodes V_i and V_j represents the presence (in G_{VU}) of a divergent path $V_i \leftarrow U_k \rightarrow V_j$ going strictly through elements of U. The presence of such bidirected edges in G represents unmeasured factors (or confounders) that may influence two variables in V; we assume that substantive knowledge permits us to decide if such confounders can be ruled out from the model. See Figure 1 for an example graph with bidirected edges.

Definition 1 (Causal-Effect Identifiability) *The* causal effect *of a set of variables T on a disjoint set of variables S is said to be* identifiable *from a graph G if the quantity $P_t(s)$ can be computed uniquely from any positive probability of the observed variables—that is, if $P_t^{M_1}(s) = P_t^{M_2}(s)$ for every pair of models M_1 and M_2 with $P^{M_1}(v) = P^{M_2}(v) > 0$ and $G(M_1) = G(M_2) = G$.*

In other words, the quantity $P_t(s)$ can be determined from the observed distribution $P(v)$ alone; the details of M are irrelevant.

The Identification of $P_x(v)$

Let X be a singleton variable. In this section we study the problem of identifying the causal effect of X on $V' = V \setminus \{X\}$, (namely, on all other variables in V), a quantity denoted by $P_x(v)$.

The easiest case

Theorem 1 *If there is no bidirected edge connected to X, then $P_x(v)$ is identifiable and is given by*

$$P_x(v) = P(v|x, pa_x)P(pa_x) \tag{6}$$

Proof: Since there is no bidirected edge connected to X, then the term $P(x|pa_x, u^x) = P(x|pa_x)$ in Eq. (4) can be moved ahead of the summation, giving

$$\begin{aligned} P(v) &= P(x|pa_x) \sum_u \prod_{\{i|V_i \neq X\}} P(v_i|pa_i, u^i)P(u) \\ &= P(x|pa_x)P_x(v). \end{aligned} \tag{7}$$

Hence,

$$P_x(v) = P(v)/P(x|pa_x) = P(v|x, pa_x)P(pa_x). \tag{8}$$

$\square$

Theorem 1 also follows from Theorem 3.2.5 of (Pearl 2000) which states that for any disjoint sets S and T in a Markovian model M, if the parents of T are measured, then $P_t(s)$ is identifiable.

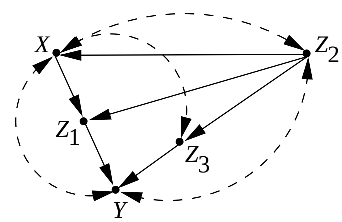


Figure 1:

A more interesting case

The case where there is no bidirected edge connected to any child of X is also easy to handle. Letting Ch_x denote the set of X's children, we have the following theorem.

Theorem 2 *If there is no bidirected edge connected to any child of X, then $P_x(v)$ is identifiable and is given by*

$$P_x(v) = \Big(\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i) \Big) \sum_x \frac{P(v)}{\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i)} \tag{9}$$

Proof: Let $S = V \setminus (Ch_x \cup \{X\})$ and $A = \prod_{\{i|V_i \in S\}} P(v_i|pa_i, u^i)$. Since there is no bidirected edge connected to any child of X, the factors corresponding to the variables in Ch_x can be moved ahead of the summation in Eqs. (4) and (5). We have

$$P(v) = (\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i)) \sum_u P(x|pa_x, u^x) \cdot A \cdot P(u), \tag{10}$$

and

$$P_x(v) = (\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i)) \sum_u A \cdot P(u). \tag{11}$$

The variable X does not appear in the factors of A, hence we augment A with the term $\sum_x P(x|pa_x, u^x) = 1$, and write

$$\begin{aligned} \sum_u A \cdot P(u) &= \sum_x \sum_u P(x|pa_x, u^x) \cdot A \cdot P(u) \\ &= \sum_x \frac{P(v)}{\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i)}. \quad \text{(by (10))} \end{aligned} \tag{12}$$

Substituting this expression into Eq. (11) leads to Eq. (9). $\square$

The usefulness of Theorem 2 can be demonstrated in the model of Figure 1. Although the diagram is quite complicated, Theorem 2 is applicable, and readily gives

$$\begin{aligned} P_x(z_1, z_2, z_3, y) &= P(z_1|x, z_2) \sum_{x'} \frac{P(x', z_1, z_2, z_3, y)}{P(z_1|x', z_2)} \\ &= P(z_1|x, z_2) \sum_{x'} P(y, z_3|x', z_1, z_2)P(x', z_2). \end{aligned} \tag{13}$$

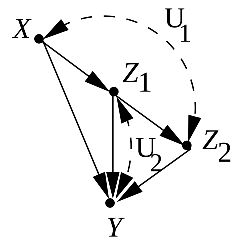


Figure 2:

The general case

When there are bidirected edges connected to the children of X, it may still be possible to identify $P_x(v)$. To illustrate, consider the graph in Figure 2, for which we have

$$P(v) = \sum_{u_1} P(x|u_1)P(z_2|z_1,u_1)P(u_1) \cdot \sum_{u_2} P(z_1|x,u_2)P(y|x,z_1,z_2,u_2)P(u_2), \quad (14)$$

and

$$P_x(v) = \sum_{u_1} P(z_2|z_1,u_1)P(u_1) \cdot \sum_{u_2} P(z_1|x,u_2)P(y|x,z_1,z_2,u_2)P(u_2). \quad (15)$$

Let

$$Q_1 = \sum_{u_1} P(x|u_1)P(z_2|z_1,u_1)P(u_1), \quad (16)$$

and

$$Q_2 = \sum_{u_2} P(z_1|x,u_2)P(y|x,z_1,z_2,u_2)P(u_2). \quad (17)$$

Eq. (14) can then be written as

$$P(v) = Q_1 \cdot Q_2, \quad (18)$$

and Eq. (15) as

$$P_x(v) = Q_2 \sum_x Q_1. \quad (19)$$

Thus, if Q_1 and Q_2 can be computed from $P(v)$, then $P_x(v)$ is identifiable and given by Eq. (19). In fact, it is enough to show that Q_1 can be computed from $P(v)$ (i.e., identifiable); Q_2 would then be given by $P(v)/Q_1$. To show that Q_1 can indeed be obtained from $P(v)$, we sum both sides of Eq. (14) over y, and get

$$P(x,z_1,z_2) = Q_1 \cdot \sum_{u_2} P(z_1|x,u_2)P(u_2). \quad (20)$$

Summing both sides of (20) over z_2, we get

$$P(x,z_1) = P(x) \sum_{u_2} P(z_1|x,u_2)P(u_2), \quad (21)$$

hence,

$$\sum_{u_2} P(z_1|x,u_2)P(u_2) = P(z_1|x). \quad (22)$$

From Eqs. (22) and (20),

$$Q_1 = P(x,z_1,z_2)/P(z_1|x) = P(z_2|x,z_1)P(x), \quad (23)$$

and from Eq. (18),

$$Q_2 = P(v)/Q_1 = P(y|x,z_1,z_2)P(z_1|x). \quad (24)$$

Finally, from Eq. (19), we obtain

$$P_x(v) = P(y|x,z_1,z_2)P(z_1|x) \sum_{x'} P(z_2|x',z_1)P(x'). \quad (25)$$

From the preceding example, we see that because the two bidirected arcs in Figure 2 do not share a common node, the set of factors (of $P(v)$) containing U_1 is disjoint of those containing U_2, and $P(v)$ can be decomposed into a product of two terms, each being a summation of products. This decomposition, to be treated next, plays an important role in the general identifiability problem.

C-components Let a path composed entirely of bidirected edges be called a *bidirected path*. The set of variables V can be partitioned into disjoint groups by assigning two variables to the same group if and only if they are connected by a bidirected path. Assume that V is thus partitioned into k groups $S_1, \ldots, S_k$, and denote by N_j the set of U variables that are parents of those variables in S_j. Clearly, the sets $N_1, \ldots, N_k$ form a partition of U. Define

$$Q_j = \sum_{n_j} \prod_{\{i|V_i \in S_j\}} P(v_i|pa_i, u^i)P(n_j), \; j = 1, \ldots, k. \quad (26)$$

The disjointness of $N_1, \ldots, N_k$ implies that $P(v)$ can be decomposed into a product of Q_j's:

$$P(v) = \prod_{j=1}^{k} Q_j. \quad (27)$$

We will call each S_j a *c-component* (abbreviating "confounded component") of V in G or a c-component of G, and Q_j the *c-factor* corresponding to the c-component S_j. For example, in the model of Figure 2, V is partitioned into the c-components $S_1 = \{X, Z_2\}$ and $S_2 = \{Z_1, Y\}$, the corresponding c-factors are given in equations (16) and (17), and $P(v)$ is decomposed into a product of c-factors as in (18).

Let $Pa(S)$ denote the union of a set S and the set of parents of S, that is, $Pa(S) = S \cup (\cup_{V_i \in S} PA_i)$. We see that Q_j is a function of $Pa(S_j)$. Moreover, each Q_j can be interpreted as the post-intervention distribution of the variables in S_j, under an intervention that sets all other variables to constants, or

$$Q_j = P_{v \setminus s_j}(s_j) \quad (28)$$

The importance of the c-factors stems from that all c-factors are identifiable, as shown in the following lemma.

Lemma 1 *Let a topological order over V be $V_1 < \ldots < V_n$, and let $V^{(i)} = \{V_1, \ldots, V_i\}$, $i = 1, \ldots, n$, and $V^{(0)} = \emptyset$. For any set C, let G_C denote the subgraph of G composed only of variables in C. Then*

(i) Each c-factor Q_j, $j = 1, \ldots, k$, is identifiable and is given by

$$Q_j = \prod_{\{i|V_i \in S_j\}} P(v_i|v^{(i-1)}). \tag{29}$$

(ii) Each factor $P(v_i|v^{(i-1)})$ can be expressed as

$$P(v_i|v^{(i-1)}) = P(v_i|pa(T_i) \setminus \{v_i\}), \tag{30}$$

where T_i is the c-component of $G_{V^{(i)}}$ that contains V_i.

Proof: We prove (i) and (ii) simultaneously by induction on the number of variables n.

Base: $n = 1$; we have one c-component $Q_1 = P(v_1)$, which is identifiable and is given by Eq. (29), and Eq. (30) is satisfied.

Hypothesis: When there are n variables, all c-factors are identifiable and are given by Eq. (29), and Eq. (30) holds for all $V_i \in V$.

Induction step: When there are $n + 1$ variables in V, assuming that V is partitioned into c-components $S_1, \ldots, S_l, S'$, with corresponding c-factors $Q_1, \ldots, Q_l, Q'$, and that $V_{n+1} \in S'$, we have

$$P(v) = Q' \prod_i Q_i. \tag{31}$$

Summing both sides of (31) over v_{n+1} leads to

$$P(v^{(n)}) = (\sum_{v_{n+1}} Q') \prod_i Q_i. \tag{32}$$

It is clear that each $S_i, i = 1, \ldots, l$, is a c-component of $G_{V^{(n)}}$. By the induction hypothesis, each $Q_i, i = 1, \ldots, l$, is identifiable and is given by Eq. (29). From Eq. (31), Q' is identifiable as well, and is given by

$$Q' = \frac{P(v)}{\prod_i Q_i} = \prod_{\{i|V_i \in S'\}} P(v_i|v^{(i-1)}), \tag{33}$$

which is clear from Eq. (29) and the chain decomposition $P(v) = \prod_i P(v_i|v^{(i-1)})$.

By the induction hypothesis, Eq. (30) holds for i from 1 to n. Next we prove that it holds for V_{n+1}. In Eq. (33), Q' is a function of $Pa(S')$, and each term $P(v_i|v^{(i-1)})$, $V_i \in S'$ and $V_i \neq V_{n+1}$, is a function of $Pa(T_i)$ by Eq. (30), where T_i is a c-component of the graph $G_{V^{(i)}}$ and therefore is a subset of S'. Hence we obtain that $P(v_{n+1}|v^{(n)})$ is a function only of $Pa(S')$ and is independent of $C = V \setminus Pa(S')$, which leads to

$$\begin{aligned}
&P(v_{n+1}|pa(S') \setminus \{v_{n+1}\}) \\
&= \sum_c P(v_{n+1}|v^{(n)}) P(c|pa(S') \setminus \{v_{n+1}\}) \\
&= P(v_{n+1}|v^{(n)}) \sum_c P(c|pa(S') \setminus \{v_{n+1}\}) \\
&= P(v_{n+1}|v^{(n)})
\end{aligned} \tag{34}$$

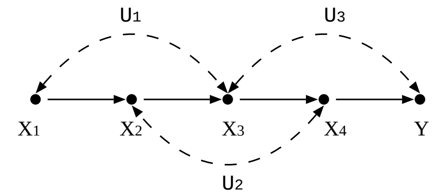


Figure 3:

□

The proposition (ii) in Lemma 1 can also be proved by using d-separation criterion (Pearl 1988) to show that V_i is independent of $V^{(i)} \setminus Pa(T_i)$ given $Pa(T_i) \setminus \{V_i\}$.

We show the use of Lemma 1 by an example shown in Figure 3, which has two c-components $S_1 = \{X_2, X_4\}$ and $S_2 = \{X_1, X_3, Y\}$. $P(v)$ decomposes into

$$P(x_1, x_2, x_3, x_4, y) = Q_1 Q_2, \tag{35}$$

where

$$Q_1 = \sum_{u_2} P(x_2|x_1, u_2) P(x_4|x_3, u_2) P(u_2), \tag{36}$$

$$\begin{aligned}
Q_2 = \sum_{u_1, u_3} & P(x_1|u_1) P(x_3|x_2, u_1, u_3) P(y|x_4, u_3) \\
& \cdot P(u_1) P(u_3).
\end{aligned} \tag{37}$$

By Lemma 1, both Q_1 and Q_2 are identifiable. The only admissible order of variables is $X_1 < X_2 < X_3 < X_4 < Y$, and Eq. (29) gives

$$Q_1 = P(x_4|x_1, x_2, x_3) P(x_2|x_1), \tag{38}$$

$$Q_2 = P(y|x_1, x_2, x_3, x_4) P(x_3|x_1, x_2) P(x_1). \tag{39}$$

We can also check that the expressions obtained in Eq.s (23) and (24) for Figure 2 satisfy Lemma 1.

The identification criterion for $P_x(v)$ Let X belong to the c-component S^X with corresponding c-factor Q^X. Let Q^X_x denote the c-factor Q^X with the term $P(x|pa_x, u^x)$ removed, that is,

$$Q^X_x = \sum_{n^X} \prod_{\{i|V_i \neq X, V_i \in S^X\}} P(v_i|pa_i, u^i) P(n^X). \tag{40}$$

We have

$$P(v) = Q^X \prod_i Q_i, \tag{41}$$

$$P_x(v) = Q^X_x \prod_i Q_i. \tag{42}$$

Since all Q_i's are identifiable, $P_x(v)$ is identifiable if and only if Q^X_x is identifiable, and we have the following theorem.

Theorem 3 *$P_x(v)$ is identifiable if and only if there is no bidirected path connecting X to any of its children. When*

$P_x(v)$ *is identifiable, it is given by*

$$P_x(v) = \frac{P(v)}{Q^X} \sum_x Q^X, \tag{43}$$

where Q^X *is the c-factor corresponding to the c-component* S^X *that contains* X.

Proof: (**if**) If there is no bidirected path connecting X to any of its children, then none of X's children is in S^X. Under this condition, removing the term $P(x|pa_x, u^x)$ from Q^X is equivalent to summing Q^X over X, and we can write

$$Q^X_x = \sum_x Q^X. \tag{44}$$

Hence from Eq.s (42) and (41), we obtain

$$P_x(v) = (\sum_x Q^X) \prod_i Q_i = (\sum_x Q^X) \frac{P(v)}{Q^X}, \tag{45}$$

which proves the identifiability of $P_x(v)$.

(**only if**) Sketch: Assuming that there is a bidirected path connecting X to a child of X, one can construct two models (by specifying all conditional probabilities) such that $P(v)$ has the same values in both models while $P_x(v)$ takes different values. The proof is lengthy and is given in (Tian & Pearl 2002). □

We demonstrate the use of Theorem 3 by identifying $P_{x_1}(x_2, x_3, x_4, y)$ in Figure 3. The graph has two c-components $S_1 = \{X_2, X_4\}$ and $S_2 = \{X_1, X_3, Y\}$, with corresponding c-factors given in (38) and (39). Since X_1 is in S_2 and its child X_2 is not in S_2, Theorem 3 ensures that $P_{x_1}(x_2, x_3, x_4, y)$ is identifiable and is given by

$$\begin{aligned} &P_{x_1}(x_2, x_3, x_4, y) = Q_1 \sum_{x_1} Q_2 \\ =&P(x_4|x_1, x_2, x_3) P(x_2|x_1) \\ &\sum_{x_1'} P(y|x_1', x_2, x_3, x_4) P(x_3|x_1', x_2) P(x_1'). \end{aligned} \tag{46}$$

A Criterion for Identifying $P_x(s)$

Let X be a singleton variable and $S \subseteq V$ be any set of variables. Clearly, whenever $P_x(v)$ is identifiable, so is $P_x(s)$. However, there are obvious cases where $P_x(v)$ is not identifiable and still $P_x(s)$ is identifiable for some subsets S of V. In this section we give a criterion for identifying $P_x(s)$.

Let $An(S)$ denote the union of a set S and the set of ancestors of the variables in S, and let $G_{An(S)}$ denote the subgraph of G composed only of variables in $An(S)$. Summing both sides of Eq. (4) over $V \setminus An(S)$, we have that the marginal distribution $P(an(S))$ decomposes exactly according to the graph $G_{An(S)}$. Therefore, if $P_x(s)$ is identifiable in $G_{An(S)}$, then it is computable from $P(an(S))$, and thus is computable from $P(v)$. A direct extension of Theorem 3 then leads to the following sufficient criterion for identifying $P_x(s)$.

Theorem 4 $P_x(s)$ *is identifiable if there is no bidirected path connecting* X *to any of its children in* $G_{An(S)}$.

When the condition in Theorem 4 is satisfied, we can compute $P_x(an(S))$ by applying Theorem 3 in $G_{An(S)}$, and $P_x(s)$ can be obtained by marginalizing over $P_x(an(S))$.

This simple criterion can classify correctly all the examples treated in the literature with X singleton, including those contrived by (Galles & Pearl 1995). In fact, for X and S being singletons, it is shown in the Appendix that if there is a bidirected path connecting X to one of its children such that every node on the path is in $An(S)$, then none of the "back-door", "front-door", and (Galles & Pearl 1995) criteria is applicable. However, this criterion is *not necessary* for identifying $P_x(s)$. Examples exist in which $P_x(s)$ is identifiable but Theorem 4 is not applicable.[7] An improved criterion that covers those cases is described in (Tian & Pearl 2002).

Conclusion

We developed new graphical criteria for identifying the causal effects of a singleton variable on a set of variables. Theorem 4 has important ramifications to the theory and practice of observational studies. It implies that the key to identifiability lies not in blocking back-door paths between X and S but, rather, in blocking back-door paths between X and its immediate successors on the pathways to S. The potential of finding and measuring intermediate variables that satisfy this condition opens new vistas in experimental design.

Acknowledgements

This research was supported in parts by grants from NSF, ONR, AFOSR, and DoD MURI program.

Appendix

In this appendix we show that Theorem 4 covers the criterion in (Galles & Pearl 1995) (which will be called the G-P criterion). The G-P criterion is for identifying $P_x(y)$ with X and Y being singletons, and it includes the "front-door" and "back-door" criteria as special cases (see (Pearl 2000, pp. 114-8)). We will prove that if there is a bidirected path connecting X to one of its children such that every node on the path is an ancestor of Y, then the G-P criterion is not applicable. There are four conditions in the G-P criterion, among which Condition 1 is a special case of Condition 3, and Condition 2 is trivial. Therefore we only need to consider Condition 3 and 4.

Proof: Assume that there is a bidirected path p from X to its child Y_1 such that every node on p is an ancestor of Y, and that there is a directed path q from Y_1 to Y. We will show by contradiction that neither Condition 3 nor Condition 4 is applicable for identifying $P_x(y)$. For any set Z, a node will be called Z-*active* if it is in Z or any of its descendants is in Z, otherwise it will be called Z-*inactive*.

(**Condition 3**) Assume that there exists a set Z that blocks all back-door paths from X to Y so that $P_x(z)$ is identifi-

[7] This implies that, contrary to claims, the criterion developed in (Galles & Pearl 1995) is *not* complete.

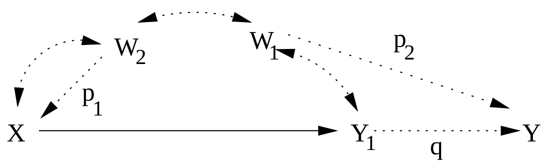
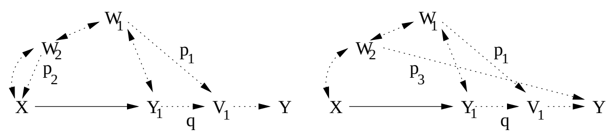


Figure 4:



(a) (b)

Figure 5:

able.[8] If every internal node on p is an ancestor of X, or if every nonancestor of X on p is Z-active, then let $W_1 = Y_1$, otherwise let W_1 be the Z-inactive non-ancestor of X that is closest to X on p (see Figure 4). If every internal node on the subpath $p(W_1, X)$ [9] is Z-active, then let $W_2 = X$, otherwise let W_2 be the Z-inactive node that is closest to W_1 on $p(W_1, X)$. From the definition of W_1 and W_2, W_2 must be an ancestor of X (or be X itself), and let p_1 be any directed path from W_2 to X. (i) If $W_1 \neq Y_1$, letting p_2 be any directed path from W_1 to Y, then from the definition of W_1 and W_2 the path $p' = (p_1(X, W_2), p(W_2, W_1), p_2(W_1, Y))$ is a back-door path from X to Y that is not blocked by Z (see Figure 4) since W_2 is Z-inactive, all internal nodes on $p(W_2, W_1)$ is Z-active, and W_1 is Z-inactive. (ii) If $W_1 = Y_1$, there are two situations:

(a) Z consists entirely of nondescendants of X. Then the path $p'' = (p_1(X, W_2), p(W_2, Y_1), q(Y_1, Y))$ is a back-door path from X to Y that is not blocked by Z.

(b) Z contains a variable Y' on $q(Y_1, Y)$ so that $P_x(z)$ is identifiable. By the definition of W_1, every node on p is an ancestor of Z. $P_x(z)$ can not be identified by Theorem 4, and the G-P criterion is not applicable for identifying $P_x(z)$ if Z contains more than one variable. If Z contains only one variable Y', then every node on p is an ancestor of Y'. If $P_x(y')$ is identifiable by Condition 3 of the G-P criterion (Condition 4 is not applicable as proved later), then from the preceding analysis there is a Y'' on the path $q(Y_1, Y')$ such that every node on p is an ancestor of Y'' and $P_x(y'')$ is identifiable. By induction, in the end we have every node on p is an ancestor of Y_1 and $P_x(y_1)$ is identifiable, which does not hold from the preceding analysis.

(**Condition 4**) Assume that there exist sets Z_1 and Z_2 that satisfy all (i)–(iv) conditions in Condition 4. Since Z_1 has to block the path $((X, Y_1), q(Y_1, Y))$, let V_1 be the variable in Z_1 that is closest to Y_1 on the path q (see Figure 5(a)). If none of the internal node on p is in $An(V_1) \setminus An(X)$ (the set of ancestors of V_1 that are not ancestors of X) or if every variable in $An(V_1) \setminus An(X)$ on p is Z_2-active, then let $W_1 = Y_1$, otherwise let W_1 be the Z_2-inactive variable in $An(V_1) \setminus An(X)$ that is closest to X on p. Let p_1 be any directed path from W_1 to V_1. If every internal node on the subpath $p(W_1, X)$ is Z_2-active, then let $W_2 = X$, otherwise let W_2 be the Z_2-inactive node that is closest to W_1 on $p(W_1, X)$. Since W_2 must be an ancestor of Y, from the definition of W_1 and W_2, there are two possible situations:

(a) W_2 is an ancestor of X or $W_2 = X$. Let p_2 be any directed path from W_2 to X (see Figure 5(a)). From the definition of W_1 and W_2, the path $p' = (p_2(X, W_2), p(W_2, W_1), p_1(W_1, V_1))$ is a back-door path from X to $V_1 \in Z_1$ that is not blocked by Z_2 that does not contain any descendant of X (see Figure 5(a)).

(b) W_2 is an ancestor of Y but not ancestor of V_1 ($W_2 \in An(Y) \setminus An(V_1)$). Let p_3 be any directed path from W_2 to Y (see Figure 5(b)). From the definition of W_1 and W_2, the path $p'' = (p_1(V_1, W_1), p(W_1, W_2), p_3(W_2, Y))$ is a back-door path from $V_1 \in Z_1$ to Y that is not blocked by Z_2 (see Figure 5(b)). $\Box$

References

Galles, D., and Pearl, J. 1995. Testing identifiability of causal effects. In Besnard, P., and Hanks, S., eds., *Uncertainty in Artificial Intelligence 11*. San Francisco: Morgan Kaufmann. 185–195.

Greenland, S.; Pearl, J.; and Robins, J. 1999. Causal diagrams for epidemiologic research. *Epidemiology* 10:37–48.

Lauritzen, S. 2000. Graphical models for causal inference. In Barndorff-Nielsen, O.; Cox, D.; and Kluppelberg, C., eds., *Complex Stochastic Systems*. London/Boca Raton: Chapman and Hall/CRC Press. chapter 2, 67–112.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligence Systems*. San Mateo, CA: Morgan Kaufmann.

Pearl, J. 1993. Comment: Graphical models, causality, and intervention. *Statistical Science* 8:266–269.

Pearl, J. 1995. Causal diagrams for experimental research. *Biometrika* 82:669–710.

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

Robins, J. 1987. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *Journal of Chronic Diseases* 40(Suppl 2):139S–161S.

Spirtes, P.; Glymour, C.; and Scheines, R. 1993. *Causation, Prediction, and Search*. New York: Springer-Verlag.

Tian, J., and Pearl, J. 2002. On the identification of causal effects. Technical Report R-290-L, Department of Computer Science, University of California, Los Angeles.

[8] A path from X to Y is said to be a *back-door* path if it contains an arrow into X.

[9] We use $p(W_1, X)$ to represent the subpath of p from W_1 to X.