Generalized Do-Calculus with Testable Causal Assumptions

Jiji Zhang
Division of Humanities and Social Sciences
California Institute of Technology
Pasadena, CA 91125
jiji@hss.caltech.edu

Abstract

A primary object of causal reasoning concerns what would happen to a system under certain interventions. Specifically, we are often interested in estimating the probability distribution of some random variables that would result from *forcing* some other variables to take certain values. The renowned *do*-calculus (Pearl 1995) gives a set of rules that govern the identification of such post-intervention probabilities in terms of (estimable) pre-intervention probabilities, assuming available a directed acyclic graph (DAG) that represents the underlying causal structure. However, a DAG causal structure is seldom fully testable given pre-intervention, observational data, since many competing DAG structures are equally compatible with the data. In this paper we extend the *do*-calculus to cover cases where the available causal information is summarized in a so-called partial ancestral graph (PAG) that represents an equivalence class of DAG structures. The causal assumptions encoded by a PAG are significantly weaker than those encoded by a full-blown DAG causal structure, and are in principle fully testable by observed conditional independence relations.

1 INTRODUCTION

For various practical purposes, such as policy analysis or decision making in general, we need to predict effects of actions or interventions before actually carrying them out. In many cases, the task is to predict what values or probability distributions of some variables would result from certain interventions on some other variables, based on data collected via passive observations without any active intervention. For such inferences, information about causal structure is needed to provide a link between pre-intervention and post-intervention probability distributions.

A prominent approach to tackling this problem makes use of graphical (or, equivalently, structural equational) representations of causal structure (Dawid 2002, Pearl 2000, Spirtes et al. 2000). Here we focus on the *do*-calculus developed by Pearl (1995, 2000). The calculus features three rules that facilitate transforming post-intervention probabilities to estimable pre-intervention probabilities. The rules are formulated with reference to a directed acyclic graph (DAG) that specifies the exact causal structure over a given set of observed and latent variables. The assumed causal diagram, however, is usually statistically indistinguishable from a lot of alternative ones, and hence is not fully testable using observational data. Sometimes such a causal diagram may be supplied by domain or expert knowledge, but when there is no substantial background causal knowledge to begin with, we may have to rely on limited causal information learnable from data.

In this paper, we give analogues of the *do*-calculus based on the formalism of ancestral graphical models. In particular, we present a *do*-calculus with respect to partial ancestral graphs (PAGs) that are fully testable given observational data. Since a PAG essentially represents an equivalence class of DAG structures (with possibly extra latent variables), it encodes substantially weaker definite causal information than a single DAG does. Yet still, it is usually possible to identify some post-intervention probabilities relative to a PAG, as we shall illustrate with the generalized *do*-calculus.

The rest of the paper is organized as follows. After introducing the necessary background in Section 2, we first develop a *do*-calculus relative to what is called maximal ancestral graphs (MAGs) in Section 3. Based on that we present a *do*-calculus relative to PAGs in Section 4. We end the paper with a simple illustration in Section 5 and a few remarks about related open questions in Section 6 .

2 PRELIMINARIES

2.1 *DO* OPERATOR AND *DO* CALCULUS

The fundamental basis for *do*-calculus and related methods is an interpretation of interventions as local and effective surgeries on a causal system. We will focus on what Pearl calls *atomic* interventions, which are what *do*-calculus deals with. Given a set of variables $\mathbf{V}$ whose causal structure can be represented by a DAG $\mathcal{G}$[1] and whose distribution factorizes according to $\mathcal{G}$, an atomic intervention forces some subset of variables $\mathbf{X}$ to take certain values $\mathbf{x}$. The intervention is supposed to be *effective* in the sense that the value of $\mathbf{X}$ is completely determined by the intervention, and *local* in the sense that the conditional distributions of other variables (variables not in $\mathbf{X}$) given their respective parents are not affected by the intervention. Symbolically the intervention will be represented by a *do*-operator: $do(\mathbf{X} = \mathbf{x})$. With these restrictions, it readily follows that:

$$P(\mathbf{V}\backslash\mathbf{X}|do(\mathbf{X} = \mathbf{x})) = \prod_{Y\in\mathbf{V}\backslash\mathbf{X}} P(Y|\mathbf{Pa}(Y)) \qquad (*)$$

Note that we use $P(\mathbf{Z}|\mathbf{W}, do(\mathbf{X} = \mathbf{x}))$ to denote a post-intervention probability distribution: the distribution of $\mathbf{Z}$ conditional on (possibly empty) $\mathbf{W}$ after the intervention $do(\mathbf{X} = \mathbf{x})$, and $\mathbf{Pa}(Y)$ to denote the set of parents of Y in the causal structure $\mathcal{G}$. Versions of the above formula can be found in Robins (1986), Pearl (2000) and Spirtes et al. (2000).

By formula $(*)$, if the pre-intervention joint probability is estimable, then any post-intervention probability is also estimable given the true causal structure. A complication comes in when there are latent variables in $\mathbf{V}$. In that case, whether a certain post-intervention probability is identifiable depends on whether, given the true causal structure, it is a functional of the pre-intervention *marginal* probability of the observed variables.

The celebrated *do*-calculus aims to deal with such situations. It gives inference rules formulated relative to a DAG with (possibly) latent variables. The antecedents of the rules involve surgeries on the causal DAG. Given a DAG $\mathcal{G}$ and a subset of variables $\mathbf{X}$ in $\mathcal{G}$, let $\mathcal{G}_{\overline{\mathbf{X}}}$ denote the graph resulting from deleting all edges in $\mathcal{G}$ that are into variables in $\mathbf{X}$, and $\mathcal{G}_{\underline{\mathbf{X}}}$ denote the graph resulting from deleting all edges in $\mathcal{G}$ that are out of variables in $\mathbf{X}$. The following proposition summarizes Pearl's *do*-calculus, which makes use of the well-known *d-separation* criterion.

Proposition 1 (Pearl). *Let $\mathcal{G}$ be the causal DAG for $\mathbf{V}$, and $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}$ be disjoint subsets of $\mathbf{V}$. The following rules are sound:*

1. *if $\mathbf{Y}$ and $\mathbf{Z}$ are d-separated by $\mathbf{X}\cup\mathbf{W}$ in $\mathcal{G}_{\overline{\mathbf{X}}}$, then*
$$P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$
2. *if $\mathbf{Y}$ and $\mathbf{Z}$ are d-separated by $\mathbf{X}\cup\mathbf{W}$ in $\mathcal{G}_{\underline{\mathbf{Z}}\overline{\mathbf{X}}}$, then*
$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W})$$
3. *if $\mathbf{Y}$ and $\mathbf{Z}$ are d-separated by $\mathbf{X}\cup\mathbf{W}$ in $\mathcal{G}_{\overline{\mathbf{X}\mathbf{Z}'}}$, then*
$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$
where $\mathbf{Z}' = \mathbf{Z}\backslash\mathbf{Ancestor}_{\mathcal{G}_{\overline{\mathbf{X}}}}(\mathbf{W})$.

The proposition follows from formula $(*)$ (see Pearl 1995 for a proof). The soundness ensures that any post-intervention probability that can be reduced via the calculus to an expression that only involves pre-intervention probabilities of observed variables is identifiable. Very recently, the completeness of the calculus was also established, in the sense that any identifiable post-intervention probability can be so reduced (Huang and Valtorta 2006, Shpister and Pearl 2006).

However, the causal DAG assumed by the *do*-calculus is not fully testable, and hence relies on substantial background knowledge. When such knowledge is not available, we may have to rely on causal information that is learnable from data. Of course such information is typically very limited, but as shown in Spirtes et al. (2000, chapter 7) and Richardson and Spirtes (2003), the limited information can still warrant interesting causal reasoning in many cases. Our aim is to devise an analogous *do*-calculus relative to partial information about the underlying causal structure, represented by a PAG.

2.2 MAGs AND PAGs

Ancestral graphs are introduced to represent data generating processes that may involve latent confounders and/or selection bias without explicitly modelling the unobserved variables (Richardson and Spirtes 2002). We are not concerned with selection bias here, so we will use only part of the machinery.

A *mixed graph* is a graph that may contain two kinds of edges: directed edges or arrows ($\rightarrow$), and bi-directed edges or double-headed arrows ($\leftrightarrow$). All the familiar

[1] The causal interpretation of a DAG is simple: an arrow from X to Y means that X is a direct cause of Y relative to $\mathbf{V}$ (Spirtes et al. 2000).

graphical notions for DAGs, *adjacency, parent/child, ancestor/descendant, path* and *directed path*, obviously remain meaningful. In addition, if there is a bi-directed edge $X \leftrightarrow Y$ in a mixed graph G, then X is called a *spouse* of Y and Y a spouse of X. An *almost directed cycle* occurs if there are two variables A and B such that A is both an ancestor and a spouse of B.

Given a path $u = \langle V_0, ..., V_n \rangle$ with $n > 1$, V_i $(1 \leq i \leq n-1)$ is a *collider* on u if the two edges incident to V_i are both into V_i, i.e., have an arrowhead at V_i; otherwise it is a *noncollider* on u. A path is called a *collider path* if every vertex on it (except for the endpoints) is a collider along the path. Let $\mathbf{L}$ be any subset of vertices in $\mathcal{G}$, an *inducing path relative to* $\mathbf{L}$ is a path on which every vertex not in $\mathbf{L}$ (except for the endpoints) is a collider on the path and every collider is an ancestor of an endpoint of the path. When $\mathbf{L}$ is empty we simply call the path an *inducing path*.[2]

A mixed graph is *ancestral* if it does not contain any directed or almost directed cycle. It is *maximal* if no inducing path is present between any two non-adjacent vertices in the graph. A *MAG* is a mixed graph that is both ancestral and maximal. Note that syntactically a DAG is a special case of MAG, simply a MAG without bi-directed edges.

There is an obvious extension of the d-separation criterion to MAGs, which, following Richardson and Spirtes (2002), we call *m-separation*.

Definition 1 (m-separation). *In a mixed graph, a path u between vertices X and Y is* **active** *(***m-connecting***) relative to a set of vertices $\mathbf{Z}$ $(X, Y \notin \mathbf{Z})$ if*

i. every non-collider on u is not a member of $\mathbf{Z}$;

ii. every collider on u is an ancestor of some member of $\mathbf{Z}$.

X and Y are said to be **m-separated** *by $\mathbf{Z}$ if there is no active path between X and Y relative to $\mathbf{Z}$.*

A nice feature of MAGs is that they can represent the marginal independence models of DAGs in the following sense: given any DAG $\mathcal{G}$ over $\mathbf{V} = \mathbf{O} \cup \mathbf{L}$ – where $\mathbf{O}$ denotes the set of observed variables, and $\mathbf{L}$ denotes the set of latent variables – there is a MAG over $\mathbf{O}$ alone such that for any three disjoint sets of variables $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \subseteq \mathbf{O}$, $\mathbf{X}$ and $\mathbf{Y}$ are d-separated by $\mathbf{C}$ in $\mathcal{G}$ if and only if they are m-separated by $\mathbf{C}$ in the MAG. The following construction gives us such a MAG:

[2] It is called a *primitive inducing path* in Richardson and Spirtes (2002).

Input: a DAG $\mathcal{G}$ over $\langle \mathbf{O}, \mathbf{L} \rangle$
Output: a MAG $\mathcal{M}_{\mathcal{G}}$ over $\mathbf{O}$

1. for each pair of variables $A, B \in \mathbf{O}$, A and B are adjacent in $\mathcal{M}_{\mathcal{G}}$ if and only if there is an inducing path between them relative to $\mathbf{L}$ in $\mathcal{G}$;

2. for each pair of adjacent variables A, B in $\mathcal{M}_{\mathcal{G}}$, orient the edge as $A \rightarrow B$ in $\mathcal{M}_{\mathcal{G}}$ if $A \in \mathbf{An}_{\mathcal{G}}(B)$; orient it as $A \leftarrow B$ in $\mathcal{M}_{\mathcal{G}}$ if $B \in \mathbf{An}_{\mathcal{G}}(A)$; orient it as $A \leftrightarrow B$ in $\mathcal{M}_{\mathcal{G}}$ otherwise.

It can be shown that $\mathcal{M}_{\mathcal{G}}$ is indeed a MAG and represents the marginal independence model over $\mathbf{O}$ (Richardson and Spirtes 2002). More importantly, notice that $\mathcal{M}_{\mathcal{G}}$ also retains the ancestral relationships — and hence causal relationships under the standard interpretation — among $\mathbf{O}$ in $\mathcal{G}$. So, if $\mathcal{G}$ is the causal DAG for $\langle \mathbf{O}, \mathbf{L} \rangle$, it is fair to call $\mathcal{M}_{\mathcal{G}}$ the **causal MAG** for $\mathbf{O}$.

Throughout the paper we will rely on the following simple example from Spirtes et al. (2000) to illustrate. Suppose we are able to measure the following random variables: *Income* (I), *Parents' smoking habits* (PSH), *Smoking* (S), *Genotype* (G) and *Lung cancer* (L) (The exact domain of each variable is not relevant for the illustration). The data, for all we know, are generated according to an underlying mechanism which might involve unmeasured common causes. Suppose, unknown to us, the structure of the causal mechanism is the one in Figure 1, where *Profession* is an unmeasured common cause of *Income* and *Smoking*.

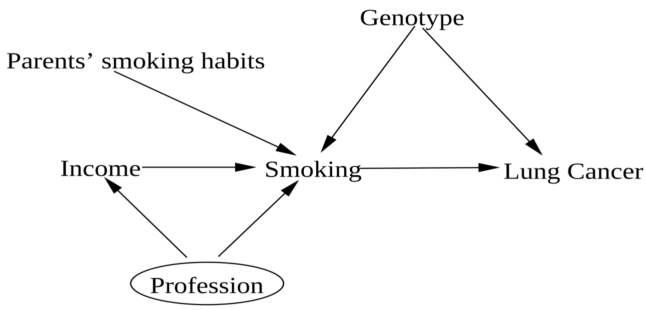

Figure 1: A causal DAG with a latent variable

This DAG structure corresponds to the MAG in Figure 2(a) (which happens to be a DAG syntactically). Different causal DAGs may correspond to the same causal MAG. So essentially a MAG represents a set of DAGs that have the exact same d-separation structure and ancestral relationships among the observed variables. A causal MAG thus carries uncertainty about what the true causal DAG is, but also represents common features shared by all possible causal DAGs.

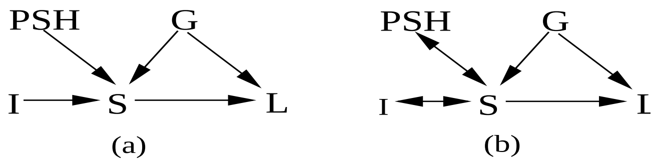

Figure 2: Two Markov Equivalent MAGs

There is then an obvious causal interpretation of MAGs, derivative from the causal interpretation of DAGs. A directed edge from A to B in a MAG means that A is a cause of B (which is a shorthand way of saying that there is a causal pathway from A to B in the underlying DAG); a bi-directed edge between A and B means that A is not a cause of B and B is not a cause of A, which implies that there is a latent common cause of A and B in the underlying DAG.

For the purpose of causal reasoning, it is important to single out a special kind of directed edges in a MAG, which we will call *visible* directed edges.

Definition 2 (**visibility**). *Given a MAG $\mathcal{M}$, a directed edge $A \to B$ in $\mathcal{M}$ is* **visible** *if there is a vertex C not adjacent to B such that there is an edge between C and A that is into A or there is a collider path between C and A that is into A and every vertex on the path is a parent of B. Otherwise $A \to B$ is said to be* **invisible**.

For example, in Figure 2(a) the edge $S \to L$ is visible, whereas the edge $I \to S$ is not. The importance of the distinction in terms of visibility is due to the following fact:

Lemma 1. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be a DAG, and $\mathcal{M}_{\mathcal{G}}$ be the MAG over $\mathbf{O}$ that represents the DAG. For any $A, B \in \mathbf{O}$, if $A \in \mathbf{Ancestor}_{\mathcal{G}}(B)$ and there is an inducing path between A and B that is into A relative to $\mathbf{L}$ in $\mathcal{G}$, then there is a directed edge $A \to B$ in $\mathcal{M}_{\mathcal{G}}$ that is invisible.*

Proof Sketch: By the construction of $\mathcal{M}_{\mathcal{G}}$, it is clear that there is a directed edge $A \to B$ in $\mathcal{M}_{\mathcal{G}}$ because $A \in \mathbf{Ancestor}_{\mathcal{G}}(B)$ and there is an inducing path between A and B relative to $\mathbf{L}$. What is left to check is that the edge $A \to B$ in $\mathcal{M}_{\mathcal{G}}$ does not satisfy the condition for visibility due to the fact that the said inducing path is into A in $\mathcal{G}$. Due to lack of space, we refer interested readers to the proof in Zhang (2006, Lemma 5.1.1). □

Lemma 1 implies that if $A \to B$ is visible in the causal MAG, then in the true causal DAG, no matter which one it is, there is no inducing path into A between A and B relative to the set of latent variables. But if a latent variable is a common cause of A and B, then there immediately is an inducing path into A via that latent common cause. Therefore, a visible directed edge between two variables implies that they do not have a latent common cause. Conversely, if a directed edge between two variables is invisible in the MAG, one can always construct a compatible DAG in which there is a latent common cause of the two variables. Thus whether a directed edge is visible conveys extremely important information about the possibility of a latent confounder.

Although a MAG encodes weaker causal assumptions than a DAG does, it is still not fully testable most of the time due to the existence of Markov equivalent alternatives. Just as different DAGs can share the exact same d-separation features and hence entail the exact same conditional independence constraints, different MAGs can entail the exact same constraints by the m-separation criterion as well. For example, the two MAGs in Figure 2 are Markov equivalent.

Several characterizations of the Markov equivalence between MAGs are available (e.g., Ali et al. 2004, Zhang and Spirtes 2005). For the present purpose, it suffices to note that all Markov equivalent MAGs have the same adjacencies and usually some common edge orientations as well. This motivates the following representation of equivalence classes of MAGs. Let *partial mixed graphs* denote such graphs that can contain four kinds of edges: $\to$, $\leftrightarrow$, $\circ\!\!-\!\!\circ$ and $\circ\!\!\to$, and hence three kinds of end marks for edges: arrowhead ($>$), tail ($-$) and circle ($\circ$).

Definition 3 (**PAG**). *Let $[\mathcal{G}]$ be the Markov equivalence class of an arbitrary MAG $\mathcal{G}$. The partial ancestral graph (PAG) for $[\mathcal{G}]$, $\mathcal{P}_{\mathcal{G}}$, is a partial mixed graph such that*

- *i. $\mathcal{P}_{\mathcal{G}}$ has the same adjacencies as $\mathcal{G}$ (and hence any member of $[\mathcal{G}]$) does;*
- *ii. A mark of arrowhead is in $\mathcal{P}_{\mathcal{G}}$ if and only if it is shared by all MAGs in $[\mathcal{G}]$; and*
- *iii. A mark of tail is in $\mathcal{P}_{\mathcal{G}}$ if and only if it is shared by all MAGs in $[\mathcal{G}]$.*[3]

Basically a PAG represents an equivalence class of MAGs by displaying all common edge marks shared by all members in the class and displaying circles for those marks that are not common, much in the same way that a Pattern (a.k.a. a PDAG or an essential graph) represents an equivalence class of DAGs. For example, Figure 3 depicts the PAG that represents the MAGs in Figure 2 (and so the DAG in Figure 1).

[3] Zhang (2006) uses the name *complete* or *maximally oriented* PAGs. Since we only consider complete ones in this paper, we will simply call them PAGs.

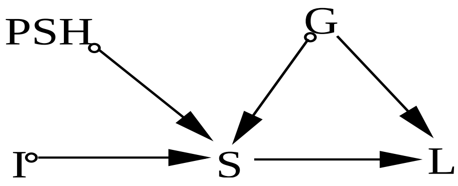

Figure 3: The PAG of the causal DAG in Figure 1

Obviously different PAGs, representing different equivalence classes of MAGs, entail different sets of conditional independence constraints. Hence a PAG is in principle fully testable by the observed conditional independence relations among the observed variables. An constraint-based algorithm to learn a PAG from an oracle of conditional independence relations is given in Zhang (2006). Score-based algorithms to learn PAGs are also under investigation.

As already mentioned, we will extend the *do*-calculus to PAGs via two steps. First we develop a *do*-calculus relative to MAGs, based on which we then develop a *do*-calculus relative to PAGs.

3 A DO-CALCULUS W.R.T. MAGs

Since a causal MAG represents an (often infinite) set of DAGs (with extra latent variables), different DAGs in the set may rule differently on whether an inference rule in Pearl's *do*-calculation is applicable. We have no intention here to contribute to the literature on conflict resolution, so we appeal to the unanimity rule. We will develop a *do*-calculus with respect to a MAG in such a way that a rule is applicable only when the corresponding rule in Pearl's calculus is applicable according to every DAG compatible with the MAG.

Since the antecedents of the rules in the *do*-calculus are formulated based on manipulations of the given DAG and in terms of d-separation, our task becomes one of defining appropriate manipulations of the given MAG and formulating appropriate rules in terms of m-separation. We first define analogous surgeries on MAGs.

Definition 4 (Manipulations of MAGs). *Given a MAG $\mathcal{M}$ and a set of variables $\mathbf{X}$ therein,*

- *the* **X-lower-manipulation** *of $\mathcal{M}$ deletes all those edges that are visible in $\mathcal{M}$ and are out of variables in $\mathbf{X}$, replaces all those edges that are out of variables in $\mathbf{X}$ but are invisible in $\mathcal{M}$ with bi-directed edges, and otherwise keeps $\mathcal{M}$ as it is. The resulting graph is denoted as $\mathcal{M}_{\underline{\mathbf{X}}}$.*
- *the* **X-upper-manipulation** *of $\mathcal{M}$ deletes all those edges in $\mathcal{M}$ that are into variables in $\mathbf{X}$, and otherwise keeps $\mathcal{M}$ as it is. The resulting graph is denoted as $\mathcal{M}_{\overline{\mathbf{X}}}$.*

We stipulate that lower-manipulation has a higher priority than upper-manipulation, so that $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ (or $\mathcal{M}_{\overline{\mathbf{X}}\underline{\mathbf{Y}}}$) denotes the graph resulting from applying the $\overline{\mathbf{X}}$-upper-manipulation to the $\mathbf{Y}$-lower-manipulated graph of $\mathcal{M}$.

A couple of comments are in order. First, unlike the case of DAGs, the lower-manipulation for MAGs may introduce new edges, i.e., replacing invisible directed edges with bi-directed edges. The reason we do this is that an invisible arrow from X to Y, as a consequence of Lemma 1, admits the possibility of a latent common cause of A and B in the underlying DAG. If so, the X-lower-manipulated DAG will correspond to a MAG in which there is a bi-directed edge between X and Y. Second, because of the possibility of introducing new bi-directed edges, we need the priority stipulation that lower-manipulation is to be done before upper-manipulation. The stipulation is not necessary for DAGs, because no new edges would be introduced in the lower-manipulation of DAGs, and hence the order does not matter.

Ideally, if $\mathcal{M}$ is the MAG that represents a DAG $\mathcal{G}$, we would like $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ to be the MAG that represents $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, where X and Y are two (possibly empty) subsets of the observed variables. But in general this is impossible, as two DAGs represented by the same MAG before a manipulation may correspond to different MAGs after the manipulation. But we still have the following:

Lemma 2. *Let $\mathcal{G}(\mathbf{O}, \mathbf{L})$ be a DAG, and $\mathcal{M}$ be the MAG of $\mathcal{G}$ over $\mathbf{O}$. Let $\mathbf{X}$ and $\mathbf{Y}$ be two possibly empty subsets of $\mathbf{O}$, and $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ be the MAG of $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. For any $A, B \in \mathbf{O}$ and $\mathbf{C} \subseteq \mathbf{O}$ that does not contain A or B, if there is an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, then there is an m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.*

Proof Sketch. The proof makes use of Lemma 1 to show essentially that for every $\mathcal{G}$ represented by $\mathcal{M}$, there is a MAG $\mathcal{M}^*$ Markov equivalent to $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ such that $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ is a subgraph of $\mathcal{M}^*$. We refer interested readers to the proof in Zhang (2006, Lemma 5.2.1). □

This lemma shows that Definition 4 is to a large extent appropriate. It implies that if an m-separation relation holds in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then it holds in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$. Hence the following corollary.

Corollary 3. *Let $\mathcal{M}$ be a MAG over $\mathbf{O}$, and $\mathbf{X}$ and $\mathbf{Y}$ be two subsets of $\mathbf{O}$. For any $A, B \in \mathbf{O}$ and $\mathbf{C} \subseteq$*

$\mathbf{O}$ *that does not contain A or B, if A and B are m-separated by $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then A and B are d-separated by $\mathbf{C}$ in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$.*

Proof Sketch: By Lemma 2, if A and B are m-separated by $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, they are also m-separated by $\mathbf{C}$ in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$, for every $\mathcal{G}$ represented by $\mathcal{M}$, which in turn implies that A and B are d-separated by $\mathbf{C}$ in $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ for every $\mathcal{G}$ represented by $\mathcal{M}$, because d-separation relations among $\mathbf{O}$ in a DAG correspond exactly to m-separation relations in its MAG. □

It is worth noting that the converse of Corollary 3 is not true in general. The reason is roughly this. Lemma 2 is true in virtue of the fact that for every $\mathcal{G}$ represented by $\mathcal{M}$, there is a MAG $\mathcal{M}^*$ Markov equivalent to $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ such that $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$ is a subgraph of $\mathcal{M}^*$. Often times there exists a $\mathcal{G}$ such that the MAG of $\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is Markov equivalent to $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. But sometimes there may not be any such DAG, and when that happens, the converse of Corollary 3 fails. For this limitation, however, Definition 4 is not to be blamed. Because no matter how we define $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, as long as it is a single graph, the converse of Corollary 3 will not hold in general. Definition 4 is already "minimal" in the following important sense: two variables are adjacent in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ if and only if there exists a DAG $\mathcal{G}$ represented by $\mathcal{M}$ such that the two variables are adjacent in $\mathcal{M}_{\mathcal{G}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}}$. In more plain terms, $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ does not have more adjacencies than necessary. We refer readers to a detailed illustration of this fact in Zhang (2006, pp. 192-193).

We are ready to present a *do*-calculus relative to a causal MAG.

Theorem 1 (***do*-calculus given a MAG**)**.** *Let $\mathcal{M}$ be the causal MAG over $\mathbf{O}$, and $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}$ be disjoint subsets of $\mathbf{O}$. The following rules are valid, in the sense that if the antecedent of the rule holds, then the consequent holds no matter which DAG represented by $\mathcal{M}$ is the true causal DAG.*

1. *if $\mathbf{Y}$ and $\mathbf{Z}$ are m-separated by $\mathbf{X} \cup \mathbf{W}$ in $\mathcal{M}_{\overline{\mathbf{X}}}$, then*

$$P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

2. *if $\mathbf{Y}$ and $\mathbf{Z}$ are m-separated by $\mathbf{X} \cup \mathbf{W}$ in $\mathcal{M}_{\underline{\mathbf{Z}}\overline{\mathbf{X}}}$, then*

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W})$$

3. *if $\mathbf{Y}$ and $\mathbf{Z}$ are m-separated by $\mathbf{X} \cup \mathbf{W}$ in $\mathcal{M}_{\overline{\mathbf{X}\mathbf{Z}'}}$, then*

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

where $\mathbf{Z}' = \mathbf{Z} \backslash \mathbf{Ancestor}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W})$.

Proof Sketch: This readily follows from Proposition 1, Corollary 3, and the fact that for every $\mathcal{G}$ represented by $\mathcal{M}$, $\mathbf{An}_{\mathcal{G}_{\overline{\mathbf{X}}}}(\mathbf{W}) \cap \mathbf{O} = \mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W})$. □

4 A DO-CALCULUS W.R.T. PAGs

The same idea applies in our formulation of a *do*-calculus relative to PAGs. Since a PAG represents an equivalence class of MAGs, we need to formulate the rules in such a way that a rule is applicable only when the corresponding rule in Theorem 1 is applicable for every MAG in the equivalence class. For this purpose, we need to establish some connections between m-connecting paths in a MAG and analogous paths in its PAG. We now proceed to define a kind of such paths in partial mixed graphs in general.

Since in general a partial mixed graph may contain ambiguous circles as well as unambiguous marks of tails and arrowheads, a path therein may contain some variables which cannot be unambiguously classified as colliders or non-colliders, as well as others that have a definite status. Let p be any path in a partial mixed graph. A (non-endpoint) vertex is called a *definite collider* on p if both incident edges are into that vertex. A (non-endpoint) vertex C is called a *definite non-collider* on p if one of the incident edges is out of C or it is $A *\!\!-\!\!\circ\, C \circ\!\!-\!\!* B$ on p such that A and B are not adjacent. Likewise, a directed edge $A \to B$ in $\mathcal{P}$ is a *definitely visible* arrow if there is a vertex C not adjacent to B such that there is an edge between C and A that is into A or there is a collider path between C and A that is into A and every vertex on the path is a parent of B. Obviously these are labelled "definite" because the available informative marks are enough to determine their respective status. Furthermore, let us call a path between A and B a *potentially directed* path from A to B if there is no arrowhead on the path pointing towards A. Variable A is called a *possible ancestor* of variable B in a partial mixed graph if there is a potentially directed path from A to B in the graph.

Definition 5 (Possibly M-Connecting Path)**.** *In a partial mixed graph, a path p between vertices A and B is* **possibly m-connecting** *relative to a set of vertices $\mathbf{Z}$ ($A, B \notin \mathbf{Z}$) if*

i. *every definite non-collider on p is not a member of $\mathbf{Z}$;*

ii. *every definite collider on p is a possible ancestor of some member of $\mathbf{Z}$.*

We then need to define relevant surgeries on PAGs. Definition 4 essentially carries over. Given a partial mixed graph $\mathcal{P}$ and a set of variables $\mathbf{X}$, $\mathcal{P}_{\overline{\mathbf{X}}}$ denotes the $\mathbf{X}$-upper-manipulated graph of $\mathcal{P}$, resulting from deleting all edges in $\mathcal{P}$ that are into variables in $\mathbf{X}$, and otherwise keeping $\mathcal{P}$ as it is. $\mathcal{P}_{\underline{\mathbf{X}}}$ denotes the $\mathbf{X}$-lower-manipulated graph of $\mathcal{P}$, resulting from deleting all definitely visible edges out of variables in $\mathbf{X}$, replacing all other edges out of vertices in $\mathbf{X}$ with bi-directed edges, and otherwise keeping $\mathcal{P}$ as it is. The priority stipulation is also the same as in Definition 4.

Given a MAG $\mathcal{M}$ and its PAG $\mathcal{P}$, since every unambiguous edge mark in $\mathcal{P}$ is also in $\mathcal{M}$ (and indeed in all MAGs equivalent to $\mathcal{M}$), it is easy to see that a m-connecting path in $\mathcal{M}$ corresponds to a possibly m-connecting path in $\mathcal{P}$. This is fortunately also true for $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ and $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, even though the latter is usually not the PAG for the former except in rare situations.

Lemma 4. *Let $\mathcal{M}$ be a MAG over $\mathbf{O}$, and $\mathcal{P}$ the PAG for $\mathcal{M}$. Let $\mathbf{X}$ and $\mathbf{Y}$ be two subsets of $\mathbf{O}$. For any $A, B \in \mathbf{O}$ and $\mathbf{C} \subseteq \mathbf{O}$ that does not contain A or B, if a path p between A and B is m-connecting given $\mathbf{C}$ in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then p, the same sequence of variables, forms a possibly m-connecting path between A and B given $\mathbf{C}$ in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$.*

Proof Sketch: It is not hard to check that for any two variables $P, Q \in \mathbf{O}$, if P and Q are adjacent in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, then they are adjacent in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. Furthermore, when they are adjacent in both $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ and $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$, every non-circle mark on the edge in $\mathcal{P}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$ is "sound" in that the mark also appears in $\mathcal{M}_{\underline{\mathbf{Y}}\overline{\mathbf{X}}}$. The lemma obviously follows. $\square$

If there is no possibly m-connecting path between A and B given $\mathbf{C}$ in a partial mixed graph, we say A and B are *definitely m-separated* by $\mathbf{C}$ in the graph. Here is the main theorem:

Theorem 2 (*do*-**calculus given a PAG).** *Let $\mathcal{P}$ be the causal PAG for $\mathbf{O}$, and $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}$ be disjoint subsets of $\mathbf{O}$. The following rules are valid:*

1. *if $\mathbf{Y}$ and $\mathbf{Z}$ are definitely m-separated by $\mathbf{X} \cup \mathbf{W}$ in $\mathcal{P}_{\overline{\mathbf{X}}}$, then*

$$P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

2. *if $\mathbf{Y}$ and $\mathbf{Z}$ are definitely m-separated by $\mathbf{X} \cup \mathbf{W}$ in $\mathcal{P}_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}$, then*

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{Z}, \mathbf{W})$$

3. *if $\mathbf{Y}$ and $\mathbf{Z}$ are definitely m-separated by $\mathbf{X} \cup \mathbf{W}$ in $\mathcal{P}_{\overline{\mathbf{X}\mathbf{Z}'}}$, then*

$$P(\mathbf{Y}|do(\mathbf{X}), do(\mathbf{Z}), \mathbf{W}) = P(\mathbf{Y}|do(\mathbf{X}), \mathbf{W})$$

where $\mathbf{Z}' = \mathbf{Z} \backslash \mathbf{PossibleAncestor}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$.

Proof Sketch: It follows from Lemma 4 and Theorem 1. The only caveat is that in general $\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W}) \neq \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$ for an arbitrary $\mathcal{M}$ represented by $\mathcal{P}$. But it is always the case that $\mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W}) \subseteq \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$, which means that $\mathbf{Z} \backslash \mathbf{An}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W}) \supseteq \mathbf{Z} \backslash \mathbf{PossibleAn}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$ for every $\mathcal{M}$ represented by $\mathcal{P}$. So it is possible that for rule (3), $\mathcal{P}_{\overline{\mathbf{X}\mathbf{Z}'}}$ leaves more edges in than necessary, but it does not affect the validity of rule (3). $\square$

5 ILLUSTRATION

Back to the case depicted in Figure 1. It is certainly impossible to fully recover this causal DAG from the data available, as the data alone by no means even indicate the relevance of the variable *Profession*. But we can, given sufficiently large sample, learn the PAG shown in Figure 3. Although the PAG reveals a limited amount of causal information, it is sufficient to identify some post-intervention quantities.

Using the *do*-calculus presented in Theorem 2, we can infer $P(L|do(S), G) = P(L|S, G)$ by rule 2, because L and S are definitely m-separated by $\{G\}$ in $\mathcal{P}_{\underline{S}}$ (Figure 4(a)); and $P(G|do(S)) = P(G)$ is true by rule 3, because G and S are definitely m-separated in $\mathcal{P}_{\overline{S}}$ (Figure 4(b)). It then follows, for instance, that

$$\begin{aligned}
P(L|do(S)) &= \sum_G P(L, G|do(S)) \\
&= \sum_G P(L|do(S), G)P(G|do(S)) \\
&= \sum_G P(L|S, G)P(G)
\end{aligned}$$

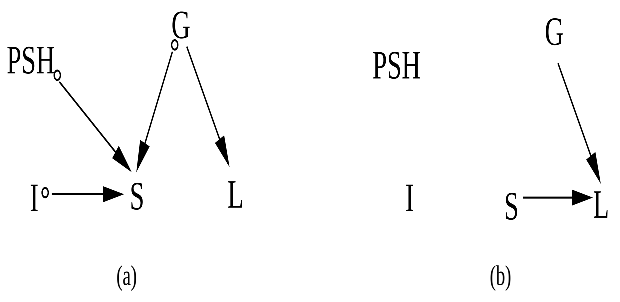

Figure 4: PAG Surgery: $\mathcal{P}_{\underline{S}}$ and $\mathcal{P}_{\overline{S}}$

By contrast, it is not valid in the *do*-calculus that $P(L|do(G), S) = P(L|G, S)$ because L and G are not definitely m-separated by $\{S\}$ in $\mathcal{P}_{\underline{G}}$, which is given in Figure 5. (Notice the bi-directed edge between L and G due to the fact that the edge $G \to L$, unlike $S \to L$, is not definitely visible in $\mathcal{P}$.)

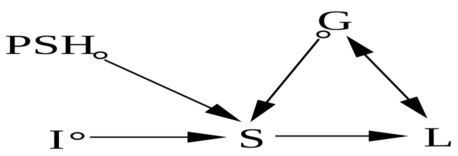

Figure 5: PAG Surgery: $\mathcal{P}_{\underline{G}}$

6 CONCLUSION

Theorem 2 gives us a *do*-calculus relative to a PAG, which is in principle learnable from data, or in other words, represents causal assumptions that are fully testable. The basic idea is that when any rule in the calculus is applicable given a PAG, the corresponding rule in Pearl's original *do*-calculus is applicable relative to each and every DAG (with possibly extra latent variables) compatible with the PAG. One open question is whether the converse is also true, i.e., whenever all DAGs compatible with the PAG sanction the application of a certain rule in the *do*-calculus, the corresponding rule is indeed applicable in our calculus relative to the PAG.

We suspect not. The rules presented in Theorem 2 are probably overly restrictive. First, even if a rule does not apply given a PAG, the corresponding rule in Theorem 1 may still apply given every MAG represented by the PAG. This may be so because a possibly m-connecting path may not actualize as m-connecting in any MAG represented by the PAG, and/or because in general $\mathbf{Ancestor}_{\mathcal{M}_{\overline{\mathbf{X}}}}(\mathbf{W})$ is a proper subset of $\mathbf{PossibleAncestor}_{\mathcal{P}_{\overline{\mathbf{X}}}}(\mathbf{W})$. Secondly, the calculus based on a MAG given in Theorem 1 may also be "incomplete", as the converse of Corollary 3 does not hold in general. Given these considerations, there may well be interesting post-intervention quantities that can be identified by Pearl's *do*-calculus given any DAG compatible with a PAG (and all these DAGs give the same answer), but cannot be identified via our *do*-calculus based on the PAG directly. We suspect, however, that such examples would be rare, if any at all. But this completeness problem is worth further investigation.

The rules in the *do*-calculus can be readily implemented, but the search for a syntactic derivation in the calculus to identify a post-intervention probability is no minor computational task. From an algorithmic point of view, *do*-calculus is probably inferior to a more algebraic method recently developed to identify intervention effects given a single causal DAG with latent variables (Tian and Pearl 2004). To adapt that approach to the case where only a PAG is given is an ongoing project.

References

Ali, R.A., T. Richardson, and P. Spirtes (2004) Markov Equivalence for Ancestral Graphs. Department of Statistics, University of Washington, Technical Report 466.

Dawid, P. (2002) Influence Diagrams for Causal Modelling and Inference. *International Statistical Review* 70: 161-189.

Huang, Y., and M. Valtorta (2006) Pearl's Calculus of Intervention Is Complete. *Proceedings of 22nd Conference on Uncertainty in Artificial Intelligence*, 217-24, Oregon: AUAI Press.

Pearl, J. (1995) Causal Diagrams for Empirical Research. *Biometrika* 82:669-710.

Pearl, J. (2000) *Causality: Models, Reasoning, and Inference.* Cambridge, UK: Cambridge University Press.

Richardson, T., and P. Spirtes (2002) Ancestral Markov Graphical Models. *Annals of Statistics* 30(4): 962-1030.

Richardson, T., and P. Spirtes (2003) Causal Inference via Ancestral Graph Models, in *Highly Structured Stochastic Systems.* Eds. P. Green, N. Hjort, and S. Richardson. Oxford University Press.

Robins, J. (1986) A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods - Applications to Control of the Healthy Worker Survivor Effect. *Mathematical Modeling* 7: 1393-1512.

Shpitser, I., and J. Pearl (2006) Identification of Conditional Interventional Distributions. *Proceedings of 22nd Conference on Uncertainty in Artificial Intelligence*, 437-44, Oregon: AUAI Press.

Spirtes, P., C. Glymour, and R. Scheines (2000) *Causation, Prediction and Search.* 2nd ed. Cambridge, MA: MIT Press.

Tian, J., and J. Pearl (2004) On the Identification of Causal Effects. Tech Report, Iowa State University, Department of Computer Science.

Zhang, J., and P.Spirtes (2005) A Transformational Characterization of Markov Equivalence for Directed Acyclic Graphs with Latent Variables, in *Proceedings of the 21th Conference on Uncertainty in Artificial Intelligence*, Oregon: AUAI Press, pp.667-674.

Zhang, J. (2006) Causal Inference and Reasoning in Causally Insufficient Systems. PhD dissertation, Department of Philosophy, Carnegie Mellon University, available at www.hss.caltech.edu/jiji/dissertation.pdf.