An Introduction to Pearl’s Do-Calculus of Intervention

Marco Valtorta
Department of Computer Science and Engineering
University of South Carolina

October 6, 2010
Contents

• Causal Bayesian networks
• The identifiability problem
• Graphical methods
• Completeness of Pearl’s do-calculus
• Summary and conclusion
Causal Bayesian Networks

• Causal Bayesian networks are Bayesian networks
  – Each variable in the graph is independent of all its non-descendents given its parents
• Causal Bayesian networks are causal
  – The directed edges in the graph represent causal influences between the corresponding variables
Key reference

Definition of Bayesian Network

Let $V$ be a finite set of finite propositional variables, $(\Omega, F, P)$ be their joint probability distribution, and $G = (V, E)$ be a dag. For each $v \in V$, let $c(v)$ be the set of all parents of $v$ and $d(v)$ be the set of all descendents of $v$. Furthermore, for $v \in V$, let $a(v)$ be $V \setminus \{d(v) \cup \{v\}\}$, i.e., the set of propositional variables in $V$ excluding $v$ and $v$'s descendents. Suppose for every subset $W \subseteq a(v)$, $W$ and $v$ are conditionally independent given $c(v)$.

Then, $C = (V, E, P)$ is called a Bayesian network [Neapolitan, 1990].
Visit to Asia Example

- Shortness of breath (dyspnoea) may be due to tuberculosis, lung cancer or bronchitis, or none of them, or more than one of them. A recent visit to Asia increases the chances of tuberculosis, while smoking is known to be a risk factor for both lung cancer and bronchitis. The results of a single chest X-ray do not discriminate between lung cancer and tuberculosis, as neither does the presence of dyspnoea [Lauritzen and Spiegelhalter, 1988].
Visit to Asia Example

\[ \alpha \text{ (Asia): } P(a) = 0.01 \]
\[ \varepsilon \text{ (or } \beta\text{): } P(e|l,t) = 1 \]
\[ P(e|l,\neg t) = 1 \]
\[ P(e|\neg l,t) = 1 \]
\[ P(e|\neg l,\neg t) = 0 \]
\[ \tau \text{ (TB): } P(t|a) = 0.05 \]
\[ P(t|\neg a) = 0.01 \]
\[ \sigma \text{ (Smoking): } P(s) = 0.5 \]
\[ \xi \text{ : } P(x|e) = 0.98 \]
\[ P(x|\neg e) = 0.05 \]
\[ \lambda \text{ (Lung cancer): } P(l|s) = 0.1 \]
\[ P(l|\neg s) = 0.01 \]
\[ \delta \text{ (Dyspnea): } P(d|e,b) = 0.9 \]
\[ P(d|e,\neg b) = 0.7 \]
\[ P(d|\neg e,b) = 0.8 \]
\[ P(d|\neg e,\neg b) = 0.1 \]

\[ \beta \text{ (Bronchitis): } P(b|s) = 0.6 \]
\[ P(b|\neg s) = 0.3 \]
Causal models are useful because they represent many different models:

- Each (hard) intervention creates a different model (“submodel” or “manipulated graph” [Spirtes et al., 2003])
- Excision semantics capture the meaning of intervention
- Intervention variables allow for soft intervention
Causality Matters!
Experts Like Causal Models
Visit to Asia Example

• Tuberculosis and lung cancer can cause shortness of breath (dyspnea) with equal likelihood. The same is true for a positive chest Xray (i.e., a positive chest Xray is also equally likely given either tuberculosis or lung cancer). Bronchitis is another cause of dyspnea. A recent visit to Asia increases the likelihood of tuberculosis, while smoking is a possible cause of both lung cancer and bronchitis [Neapolitan, 1990].
A Simple Smoking and Lung Cancer Model

- The simplest model
- \( X: \) Smoking, \( Y: \) Lung cancer
- \( X \) and \( Y \) are observable
Fisher’s Smoking and Lung Cancer Model

- Smoking-lung cancer model with genotype
- Genotype (U) is unobservable
Smoking and Lung Cancer Model with an Intermediate Variable

- Smoking and lung cancer model with genotype and tar deposits in the lungs
- Tar deposits (Z) are observable

\[ U \]
\[ X \rightarrow Z \rightarrow Y \]
Cochran’s Soil Fumigants Example

**X**: soil fumigants  
**Y**: oat crop yield  
**Z₀**: last year’s eelworm population (unknown)  
**Z₁**: eelworm populations before treatments  
**Z₂**: eelworm populations after treatments  
**Z₃**: eelworm population at the end of the season  
**B**: population of birds (unknown)

We wish to assess the total effect of the fumigants on crop yield. But controlled randomized experiment are unfeasible, and **Z₀** and **B** are unknown.
The Excision Semantics for Intervention

• The effect of an intervention that fixes the values of a (subset of the) variable(s) $X$ to $x$ is to remove the links getting into $X$ and force $X$ to have the fixed value(s) $x$

• The deletion of the link(s) into $X$ represents the understanding that, whatever relationship existed between $X$ and its parents prior to the action, that relationship is no longer in effect after we perform the action
Effect of Interventions

\[ P(x_1, \ldots, x_n \mid \hat{x}_i') = \begin{cases} \prod_{j \neq i} p(x_j \mid pa_j) & \text{if } x_i = x_i' \\ 0 & \text{if } x_i \neq x_i' \end{cases} \]  
(3.10)

\[ P(x_1, \ldots, x_n \mid \hat{x}_i') = \begin{cases} \frac{P(x_1, \ldots, x_n)}{P(x_i' \mid pa_i)} & \text{if } x_i = x_i' \\ 0 & \text{if } x_i \neq x_i' \end{cases} \]  
(3.11)

\[ P(x_1, \ldots, x_n \mid \hat{x}_i') = \begin{cases} P(x_1, \ldots, x_n \mid x_i', pa_i)P(pa_i) & \text{if } x_i = x_i' \\ 0 & \text{if } x_i \neq x_i' \end{cases} \]  
(3.12)

• Theorem 3.2.2 (adjustment for direct causes)

Let PA_i denote the set of direct causes of variable X_i, and let Y be any set of variables disjoint of \( \{X_i \cup PA_i\} \). The effect of the intervention \( do(X_i = x_i') \) on Y is given by

\[ P(y \mid \hat{x}_i') = \sum_{pa_i} P(y \mid x_i', pa_i)P(pa_i) \]  
(3.13),

where \( P(y \mid x_i', pa_i) \) and \( P(pa_i) \) represent pre-intervention probabilities.
Imperfect Interventions
The effect of an atomic intervention $\text{do}(X_i=x'_i)$ is encoded by adding to $G$ a link $F_i \rightarrow X_i$, where $F_i$ is a new variable taking values in $\{\text{do}(x'_i), \text{idle}\}$, $x'_i$ ranges over the domain of $X_i$, and "idle" represents no intervention. Then we define:

$P'(x_i \mid p_{a'}) = \begin{cases} P(x_i \mid p_{a_i}) & \text{if } F_i = \text{idle} \\ 0 & \text{if } F_i = \text{do}(x'_i) \text{ and } x_i = x'_i \\ 1 & \text{if } F_i = \text{do}(x'_i) \text{ and } x_i \neq x'_i \end{cases}$

$P(x_1, \ldots, x_n \mid \hat{x}') = P'(x_1, \ldots, x_n \mid F_i = \text{do}(x'_i))$
Contents

• Causal Bayesian networks

• The identifiability problem

• Graphical methods

• Completeness of Pearl’s do-calculus

• Summary and conclusion
Identifiable Quantities

• The quantity $P_t(s)$ is identifiable if, given the causal Bayesian network $G$, the quantity $P_t(s)$ can be determined from the distribution of the observed variables $P(n)$ alone
  
  — Alternative notation:
  
  $P(s \mid \text{set}(t)), P(s \mid \text{do}(t)), P(s \mid t), P(s \mid| t)$

• For example, we may want to assess the total effect of the fumigants ($t=\{X\}$) on yield ($s=\{Y\}$), denoted $P_t(s)$, when we have the graph and $P(X,Z_1,Z_2,Z_3,Y)$
Unidentifiability Example (1)

- All the variables are binary.
- $P(U=0) = 0.5$,
- $P(X=0 | U) = (0.6, 0.4),$
- $P(Y=0 | X, U) = \begin{array}{c|cc}
    \text{Y=0} & \text{X=0} & \text{X=1} \\
    \hline
    U=0 & 0.7 & 0.2 \\
    U=1 & 0.2 & 0.7 \\
\end{array}$
Unidentifiability Example (2)

• Note that

\[ P(X, Y) = \sum_U P(Y \mid X, U)P(X \mid U)P(U) \]

• We get:

<table>
<thead>
<tr>
<th></th>
<th>X =0</th>
<th>X =1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y =0</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(=0.7 \times 0.6 \times 0.5 + 0.2 \times 0.4 \times 0.5)</td>
<td>0.25</td>
</tr>
<tr>
<td>Y =1</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

• Because of the excision semantics, the link from U to X is removed, and we have:

\[ P_X(Y) = \sum_U P(Y \mid X, U)P(U) \]

• So, \( P_{X=0}(Y=0) = (0.7 \times 0.5) + (0.2 \times 0.5) = 0.45 \)
Unidentifiability Example (3)

- All the variables are still binary.
- $P(U=0) = 0.5$
- $P(X=0|U) = (0.7, 0.3)$
- $P(Y=0|X,U) = \begin{array}{c|cc}
    Y=0 & X=0 & X=1 \\
    \hline
    U=0 & 0.65 & 0.15 \\
    U=1 & 0.15 & 0.65
\end{array}$
Unidentifiability Example (4)

- Using

\[ P(X, Y) = \sum_U P(Y \mid X, U) P(X \mid U) P(U) \]

- We still get:

<table>
<thead>
<tr>
<th></th>
<th>X =0</th>
<th>X = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y =0</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Y=1</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

- From

\[ P_X(Y) = \sum_U P(Y \mid X, U) P(U) \]

- We have \( P_{X=0}(Y=0) = (0.65 \times 0.5) + (0.35 \times 0.5) = 0.4 < 0.45 \)
- So, \( P_X(Y) \) is unidentifiable in this model
Identifiability Problem

• For a given causal Bayesian network, decide whether \( P_t(s) \) is identifiable or not

• If \( P_t(s) \) is identifiable, give a closed-form expression for the value of \( P_t(s) \) in terms of distributions derived from the joint distribution of all observed quantities, \( P(n) \)
Contents

• Causal Bayesian networks
• The identifiability problem
• Graphical methods
• Completeness of Pearl’s do-calculus
• Summary and conclusion
Graphical Methods

- Back-door criterion
- Front-door criterion
- The three “do-calculus” inference rules
- Causal effect with single intervention variable
- Extensions of back-door criterion
- Extensions of front-door criterion
Back-Door Criterion

• A set of variables Z satisfies the back-door criterion relative to an ordered pair of variables \((X_i, X_j)\) in a DAG G if
  
  (1) no node in Z is a descendant of \(X_i\), and
  
  (2) Z blocks every path between \(X_i\) and \(X_j\) that contains a directed link into \(X_i\)

• Similarly, if \(X\) and \(Y\) are two disjoint subsets of nodes in G, then Z is said to satisfy the back-door criterion relative to \((X,Y)\) if it satisfies the criterion relative to any pair \((X_i, X_j)\) such that \(X_i \in X, X_j \in Y\)
Use of the Back-Door Criterion

• If a set of variables $Z$ satisfies the back-door criterion relative to $(X, Y)$, then the causal effect of $X$ on $Y$ is identifiable and is given by the formula:

$$P_x(y) = \sum_z P(y \mid x, z)P(z)$$
Using the Back-Door Criterion: an Example

- We can use the back-door criterion on the simple lung cancer model
- Here $Z = \emptyset$, and we get:

$$P_x(y) = P(y \mid x)$$

\[ \begin{align*}
 &\quad X \quad \rightarrow \quad Y \\
\end{align*} \]
Smoking and the genotype theory

- X: Smoking
- Y: Lung Cancer
- Z: Tar Deposits
- Can we get $P(y | \hat{x})$?
Effect of Smoking on Lung Cancer in the Presence of Tar Deposits

• We compute \( P(y \mid \tilde{x}) \) in two steps: the causal effect of \( x \) on \( z \) and the causal effect of \( z \) on \( y \).

• By application of the backdoor criterion with an empty set of concomitants, \( P(z \mid \tilde{x}) = P(z \mid x) \)

• And, by the backdoor criterion with \( X \) as a concomitant, \( P(y \mid \tilde{z}) = \sum_x P(y \mid x, z)P(x \mid z) \)

• Putting things together: \( P(y \mid \tilde{x}) = \sum_z P(z \mid x) \sum_{x'} P(y \mid x', z)P(x') \)
Front-Door Criterion

- A set of variables Z is said to satisfy the front-door criterion relative to an ordered pair of variables (X, Y) if
  1. Z intercepts all directed paths from X to Y
  2. there is no back-door path from X to Z
  3. all back-door paths from Z to Y are closed by X
Use of the Front-Door Criterion

- If \( Z \) satisfies the front-door criterion relative to \((X,Y)\) and if \( P(x,z) > 0 \), then the causal effect of \( X \) on \( Y \) is identifiable and is given by formula:

\[
 P_x(y) = \sum_z P(z \mid x) \sum_{x'} P(y \mid x', z) P(x')
\]

- The front-door criterion may be obtained by a double application of the back-door criterion, as shown in the tar deposits example.
Using the Front-Door Criterion: an Example

• In the smoking-lung cancer model with genotype and tar, we may use the front-door criterion and get:

\[
P_X(y) = \sum_z P(z \mid x) \sum_{x'} P(y \mid x', z) P(x')
\]
Notation for Graphs with Link Deletion

- In a causal Bayesian network $G$, let $X, Y, Z$ be arbitrary disjoint set of nodes.
- $G_{\bar{X}}$: the graph obtained by deleting from $G$ all directed links pointing to nodes in $X$.
- $G_X$: the graph obtained by deleting from $G$ all directed links emerging from nodes in $X$.
- $G_{\bar{X}Z}$: the deleting of both incoming and outgoing links.
Pearl’s Calculus of Interventions

• For any disjoint subsets of variables X, Y, Z, and W:

  • Rule 1 (Insertion/Deletion of Observations):
    
    \[ P_x(y|z,w) = P_x(y|w) \text{ if } (Y \perp Z|X,W)_{G_X} \]

  • Rule 2 (Action/Observation Exchange):
    
    \[ P_{x,z}(y|w) = P_x(y|z,w) \text{ if } (Y \perp Z|X,W)_{G_{XZ}} \]

  • Rule 3 (Insertion/Deletion of Actions):
    
    \[ P_{x,z}(y|w) = P_x(y|w) \text{ if } (Y \perp Z|X,W)_{G_{X,Z(w)}} \]
    
    – Where Z(W) is the set of Z-nodes that are not ancestors of any W-nodes in \( G_X \)
To determine the causal effect of Selenium on Keshan Disease, we need to find a set $Z$ of variables (set of concomitants) that satisfies the back-door criterion.

$Z=\{\text{Region of China}\}$ is an answer.

$Z=\{\text{Genotype}\}$ would also be an answer if it were observable.
The Smoking Example

Based on rule 2, we have $P(z \mid \hat{x}) = P(z \mid x)$
The Smoking Example

2. \[ P(y \mid \hat{z}) = \sum_x P(y \mid x, \hat{z}) P(x \mid \hat{z}). \]

\[ P(x \mid \hat{z}) = P(x) \quad \text{if} \quad (Z \perp X)_{G_{\hat{Z}}} \quad \text{rule 3} \]

\[ P(y \mid x, \hat{z}) = P(y \mid x, z) \quad \text{if} \quad (Z \perp Y \mid X)_{G_{\hat{Z}}} \quad \text{rule 2} \]

So, \[ P(y \mid \hat{z}) = \sum_x P(y \mid x, z) P(x) \]

Note: we can use the same process to prove the back-door formula
The Smoking Example

Based on 1, 2, and above, we get:

\[ P(y \mid \hat{x}) = \sum_z P(y \mid z, \hat{x})P(z \mid \hat{x}), \]

\[ P(y \mid z, \hat{x}) = P(y \mid \hat{z}, \hat{x}), \quad \text{rule 2} \]

\[ P(y \mid z, \hat{x}) = P(y \mid \hat{z}), \quad \text{rule 3} \]

Based on 1, 2, and above, we get:

\[ P(y \mid \hat{x}) = \sum_z P(z \mid x) \sum_{x'} P(y \mid x', z)P(x') \]

Note: this is also a proof of the front-door formula
The Smoking Example

- 4 \[ P(y, z \mid \hat{x}) = P(y \mid z, \hat{x})P(z \mid \hat{x}). \]

See 1 and the third formula in 3; we have:

\[ P(y, z \mid \hat{x}) = P(y \mid \hat{z})P(z \mid x) = P(z \mid x)\sum_{x'} P(y \mid x', z)P(x') \]

- 5 \[ P(x, y \mid \hat{z}) = P(y \mid x, \hat{z})P(x \mid \hat{z}) = P(y \mid x, z)P(x). \]

See 2
Figure 3.7: (a) A bow pattern: a confounding arc embracing a causal link $X \rightarrow Y$, thus preventing the identification of $P(y|x)$ even in the presence of an instrumental variable $Z$, as in (b). (c) A bowless graph that still prohibits the identification of $P(y|x)$. 
Some Identifiable Models

Figure 3.8: Typical models in which the effect of $X$ on $Y$ is identifiable. Dashed arcs represent confounding paths, and $Z$ represents observed covariates.
Why They are Identifiable

• (a), (b) rule 2,
• (c), (d) back-door
• (e) front-door
• (f) \[ P(y | \hat{x}) = \sum_{z_1, z_2} P(y | z_1, z_2, \hat{x})P(z_1, z_2 | \hat{x}) \]
  \[ P(y | z_1, z_2, \hat{x}) = P(y | z_1, z_2, x) \quad \text{rule 2} \]
  \[ P(y | \hat{x}) = \sum_{z_1, z_2} P(y | z_1, z_2, x)P(z_1 | x) \sum_{x'} P(z_2 | z_1, x')P(x'). \]
• (g) \[ P(y | \hat{x}) = \sum_{z_1, z_2} \sum_{x'} P(y | z_1, z_2, x')P(x' | z_2)P(z_1 | x, z_2)P(z_2) \]
Why g?

- Z1 block all directed paths from X to Y
- Z2 blocks all back-door paths between Y and Z1 in $G_{\bar{X}}$
- Putting the pieces together, we obtain the claimed result

$$P(y | \hat{x}) = \sum_{z_1, z_2} P(y | \hat{x}, z_1, z_2)P(z_1, z_2 | \hat{x})$$

$$P(y | \hat{x}, z_1, z_2) = P(y | \hat{x}, \hat{z}_1, z_2) =$$

$$P(y | \hat{z}_1, z_2) = P(y | \hat{z}_1, z_2) =$$

$$P(y | x', \hat{z}_1, z_2) = \sum_{x'} P(y | x', \hat{z}_1, z_2)P(x' | \hat{z}_1, z_2)$$

$$P(y | x', z_1, z_2) = P(y | x', z_1, z_2)$$

$$P(x' | \hat{z}_1, z_2) = P(x' | z_2)$$

$$P(z_1, z_2 | \hat{x}) = P(z_2 | \hat{x})P(z_1 | \hat{x}, z_2) =$$

$$P(z_2)P(z_1 | \hat{x}, z_2) = P(z_2)P(z_1 | x, z_2)$$
Several More Nonidentifiable Models

Figure 3.9: Typical models in which $P(y|x)$ is not identifiable.
Derived Algorithms and Criteria

• The back-door and front-door criteria can be obtained by using the three inference rules

• The algorithm and derived criteria listed below can also be obtained by using the three rules
  – Causal effect with single intervention variable (Galles & Pearl)
  – Extension of Back-door criterion (Robins & Pearl)
  – Extension of Front-door criterion (Kuroki & Miyakawa)

• For the causal effect of X and Y, where X and Y are individual variable, Tian proved a beautifully simple necessary and sufficient criterion [Thm.3.6.1 Pearl 2009]
Contents

• Causal Bayesian networks
• The identifiability problem
• Graphical methods
• Completeness of Pearl’s do-calculus
• Summary and Conclusion
Pearl’s Conjecture

- Pearl conjectured that $P_t(s)$ is identifiable if and only if all terms involving intervention can be replaced by terms involving only observations, by successive application of the three inference rules of the “do calculus”
Completeness of the do-calculus

- Various papers by Tian, Shpitser, Pearl, Huang, Valtorta
- The dates are backwards because of publication delays
Pearl’s Graphical Inference Rules Are Complete

• Exploiting the complete Identify algorithm on normal causal Bayesian networks, we prove that Pearl’s three inference rules are complete

• This result confirms Pearl’s conjecture
Outline of the Proof

• The sound and complete algorithm for computing $P_T(S)$ is obtained by using lemma 1 and lemma 2

• We show that these two lemmas can be obtained by just using the three inference rules and standard probability manipulations. Therefore, the sufficiency of the three rules is proved

Three inference rules $\rightarrow$ Tian & Pearl’s two lemmas $\rightarrow$ A sound & complete algorithm
Contents

• Causal Bayesian networks
• The identifiability problem
• Graphical methods
• Completeness of Pearl’s do-calculus

• Summary and Conclusion
Summary: Algorithms for the Identifiability Problem in Causal Bayesian Networks

- Each variable is independent of all its non-descendents given its direct parents.
- The directed edges in the graph represent causal influences between the corresponding variables.
- Some variables (e.g., genotype) are not observable.
- Can we identify a causal effect from observable data alone?
  - E.g., can we compute the probability of lung cancer given smoking in the three models of the left, even when the genotype is not observable?
- Are there efficient, sound, and complete algorithms to solve the above identifiability problem?
Slides available at
http://www.cse.sc.edu/~mgv/talks/AIM2010.ppt