Pearl, Causality, 2nd Edition – Chapter 4.2-4.3: Actions, Plans and Direct Effects

Claudio López

4.2 Acciones condicionales y políticas estocásticas.

Las intervenciones pueden involucrar políticas complejas en las que una variable X se hace para responder de una manera específica a un conjunto Z de otras variables.

$$x=g(z) \rightarrow P^*(x/z)$$

Identificar el efecto de tales políticas es equivalente a identificar la expresión $P(y/do(x),z)$

Ejemplo:

Sea $P(y/do(X=g(z))$. Para ser computado, condicionamos Z y escribimos:

$$P(y/do(X=g(z)))=\sum_z P(y/do(X=g(z)),z)P(z/do(X=g(z)))$$

$$=\sum_z P(y/do(x),z)|_{x=g(z)}P(z)$$

$$=E_z[P(y/do(x),z)|_{x=g(z)}]$$

Es igual $P(z/do(X=g(z)) = P(z)$.

Z no puede ser un descendiente de X.

4.2 Acciones condicionales y políticas estocásticas.

El efecto causal de una política do(X=g(z)) puede evaluarse directamente a partir de la expresión de P (y/do(x),z) simplemente sustituyendo g(z) por x y tomando la expectativa sobre Z (utilizando el Distribución P(z)).

Si es identificable X=g(z), entonces X=x también es identificable, ya que podemos tener g(z)=x. Pero al contrario no, ya que el condicionamiento en Z podría crear dependencias impidiendo la reducción de P(y/do(x),z) a una expresión sin sombrero.

Se considera la intervención estocástica como un proceso de azar donde la intervención incondicional X=x se aplica con probabilidad P*(x/z). Por lo tanto, dado Z=z, la intervención do(X=x) ocurrirá con la probabilidad P*(x/z), y producirá un efecto causal dado por P(y/do(x),z).

4.2 Acciones condicionales y políticas estocásticas.

Promediando sobre x y z da el efecto (en Y) de una política estocástica P*(x/z):

$$P(y)|_{P^*(x/z)} = \Sigma_x \Sigma_z P(y/do(x),z)P^*(x/z)P(z)$$

Debido a que P*(x / z) se especifica externamente, vemos nuevamente que la identificabilidad de P(y/do(x),z) es una condición necesaria y suficiente para la identificación de cualquier política estocástica que da forma a la distribución de X por el resultado de Z.

4.2 Acciones condicionales y políticas estocásticas.

Por OTRA parte, la planificación de acciones STRIPS-like, cuyo efecto inmediato X=x depende de la satisfacción de algunas precondiciones habilitantes C(w) de un conjunto de variables W. Para representar esta acción, establecemos que $Z = W \cup PA_X$ y conjunto

$$P^*(x \mid z) = \begin{cases} P(x \mid pa_X) & \text{if } C(w) = \text{false}, \\ 1 & \text{if } C(w) = \text{true and } X = x, \\ 0 & \text{if } C(w) = \text{true and } X \neq x. \end{cases}$$

4.3 ¿Cuándo es identificable el efecto de una acción?

En el cap. 3 vimos varios criterios gráficos para identificar el efecto que causa una variable en otra P(y/do(x)). El criterio de la puerta trasera y la puerta delantera son casos especiales de una clase mas general de modelos Semi-Markovianos.

La Fig. 3.1 y 3.8.f son casos donde el criterio de la puerta trasera o delantera no son suficientes para identificar P(y/x) pero aun así reducible y por lo tanto identificable.

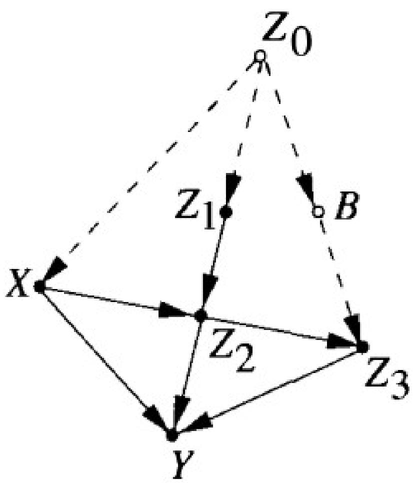
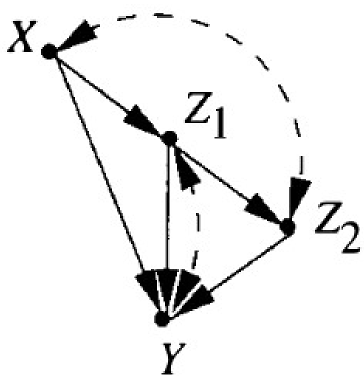


Figure 3.1 A causal diagram representing the effect of fumigants (X) on yields (Y).



(f)

4.3.1 Condiciones graficas para la identificación.

El teorema 4.3.1 Establece 4 condiciones graficas para identificar P(y/x) cuando X y Y son nodos singleton en la grafica. El teorema 4.3.2 luego afirma la integridad (o necesidad) de estas cuatro condiciones.

Aun es un tema abierto según Pearls.

Teorema 4.3.1: Sea X y Y dos variables singleton en un modelo semi-Markoviano caracterizado por el gráfico G. Una condición suficiente para la identificabilidad de P(y/x) es que G cumple una de las siguientes cuatro condiciones.

Teorema 4.3.1

1. No hay una puerta trasera de X a Y en G.
$(X \perp\!\!\!\perp Y)_{G_{\overline{X}}}$

2. No hay un camino dirigido de X a Y en G.

3. Existe un conjunto de nodos B que bloquea todas las rutas de la puerta trasera de X a Y para que la P(b/do(x)) sea identificable. (Un caso especial de esta condición ocurre cuando B consiste enteramente en no descendientes de X, en cuyo caso el P(b/do(x)) se reduce inmediatamente a P(b).

Teorema 4.3.1

4. Existen conjuntos de nodos Z1 y Z2 tales que:

i) Z1 bloquea cada ruta dirigida de X a Y. $(X \perp\!\!\!\perp Y|Z1)_{G_{\overline{Z1}\,\underline{X}}}$

ii) Z2 bloquea todos los caminos de puerta trasera entre Z1 y Y. $(Y \perp\!\!\!\perp Z1|Z2)_{G_{\overline{X\,Z1}}}$

iii) Z2 bloquea todos los caminos de puerta trasera entre X y Z1. $(X \perp\!\!\!\perp Y1|Z2)_{G_{\overline{X}}}$

iv) Z2 no activa ninguna ruta de puerta trasera de X a Y. $(X \perp\!\!\!\perp Y|Z1,Z2)_{\overline{Z1\,X\,Z2}}$ (Esta condición se mantiene si (i) - (iii) se cumplen y ningún miembro Z_jX (Z1) de Z2 es un descendiente de X.)

Pruebas del teorema 3.4.1

Condición 1: Esta condición se deriva directamente de la Regla 2 (Teorema 3.4.1). Si **$(Y \perp\!\!\!\perp X)_{G_x}$** entonces podemos cambiar inmediatamente **P (y/do(x)) a P (y/x)**, por lo que la consulta es identificable.

Condición 2: Si no hay una ruta dirigida de **X** a **Y** en **G**, entonces **$(Y \perp\!\!\!\perp X)_{G_x}$** Por lo tanto, según la Regla 3, **P(y/do(x)) = P(y)** y por lo tanto la consulta es identificable.

Condición 3: Si hay un conjunto de nodos **B** que bloquean todas las rutas de puerta trasera de **X** a **Y** (es decir, **$(Y \perp\!\!\!\perp X|B)$**), entonces podemos expandir

P(y/do(x)) como **Σ P(y/do(x),b)P(b/do(x))** y, según la Regla 2, se reescribe **P(y/do(x), b)** como **P(y/x, b)**. Si la consulta **(b/do(x))** es identificable, entonces la consulta original también debe ser identificable. Vea los ejemplos en la Figura 4.1.

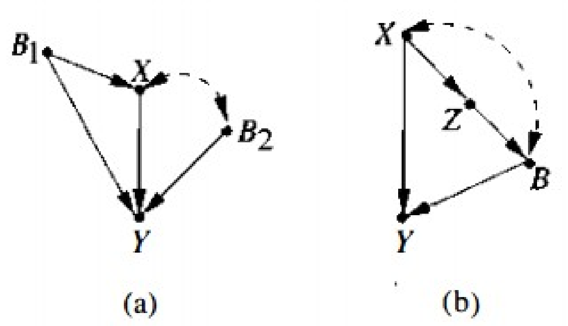


Figure 4.1 Condition 3 of Theorem 4.3.1. In (a), the set $\{B_1, B_2\}$ blocks all back-door paths from X to Y, and $P(b_1, b_2 \mid \hat{x}) = P(b_1, b_2)$. In (b), the node B blocks all back-door paths from X to Y, and $P(b \mid \hat{x})$ is identifiable using Condition 4.

Prueba del teorema 4.3.1

Condición 4. Si hay un conjunto de nodos **Z1** que bloquean todos los caminos dirigidos de X a Y y un conjunto de nodos **Z2** que bloquean todos los caminos de puerta trasera entre **Y** y **Z1** en $\mathbf{G_{\overline{x}}}$, entonces expandimos

P(y/do(x)) = Σ P(y/do(x),z1,z2)P(z1,z2|do(x)) y reescribe **P(y/do(x),z1,z2)** como": **P(y/do(x),z1^,z2)** usando la Regla 2, ya que todas las rutas de puerta trasera entre **Z1** e **Y** están bloqueadas por **Z2** en $\mathbf{G_{\overline{x}}}$**.** Podemos reducir **P(y/do(x),do(z1),z2)** a **P(y/do(z1),z2)** usando la Regla 3, ya que $\mathbf{(Y \perp\!\!\!\perp X \mid Z1,Z2)_{G_{\overline{x}\, \overline{z1}\, \overline{z2}}}}$. Podemos reescribir **P(y/do(z1),z2)** como **P(y/z1,z2)** si $\mathbf{(Y \perp\!\!\!\perp Z1 \mid Z2)_{G_{\underline{z1}}}}$.

Teorema 4.2.1

Las cuatro condiciones del Teorema 4.3.1 son necesarias para la identificación en el cálculo do. Esto es, si las cuatro condiciones del Teorema 4.3.1 fallan en un gráfico G, entonces no existe una secuencia de reglas de inferencia que reduzca P(y/do(x)) a una expresión hat-feed.

4.3.2 Observaciones sobre la eficiencia

Al ver el teorema 4.3.1 de puede observar que las condiciones 1 y 3 son búsquedas exhaustivas. Lo cual solo es refutable al demostrar que no existe dicho conjunto de bloqueo B.

Teorema 4.3.3: Si **P(b_i/do(x))** es identificable para un conjunto mínimo B_i, entonces **P(b_j/do(x))** es identificable para cualquier otro conjunto mínimo **B_j**.

El teorema 4.3.3 nos permite probar la condición 3 con un único conjunto de bloqueo mínimo B. Si cumple con los requisitos de la condición 3, entonces la consulta es identificable; de lo contrario, la condición 3 no puede ser satisfecha.

Lema 4.3.4: Si la consulta **P(y/do(x))** es identificable y si un conjunto de nodos **Z** se encuentra en una ruta dirigida de **X** a **Y**, entonces la consulta **P(z/do(x))** es identificable.

Teorema 4.3.5

Sean **Y1** e **Y2** dos subgrupos de nodos, de manera que (i) no hay nodos **Y1** son descendientes de **X** o (ii) todos los nodos **Y1** e **Y2** son descendientes de **X** y todos los nodos **Y1** no son descendientes de **Y2**. Existe una secuencia reductora para **P(y1,y2/do(x))** (según el Corolario 3.4.2) si y solo si hay secuencias reductoras para **P(y1/do(x))** y **P(Y2/do(x),y1**).

El teorema 4.3.5 garantiza que, si hay una secuencia reductora para P(y1,y2/do(x)), entonces siempre deberíamos poder encontrar dicha secuencia tanto para P(y1/do(x)) como para P(y2/do(x),y2) por elección adecuada de Y1.

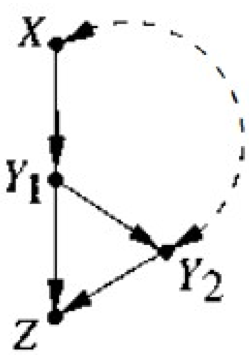


Figure 4.3 Theorem 4.3.1 ensures a reducing sequence for $P(y_2 \mid \hat{x}, y_1)$ and $P(y_1 \mid \hat{x})$, although none exists for $P(y_1 \mid \hat{x}, y_2)$.

Teorema 4.3.6

Si existe un conjunto Z1 que cumple con todos los requisitos en la Condición 4, entonces el conjunto que consiste en los hijos de X interceptados con los antepasados de Y también cumplirá con todos los requisitos para Z1 en la Condición 4.

El teorema 4.3.6 elimina la necesidad de buscar Z1 en la Condición 4 del Teorema 4.3.1.

4.3.3 Derivando una expresión de forma cerrada para consultas de control

El teorema 4.3.1 ofrece un algoritmo para determinar la capacidad de identificación de una consulta de control, además, proporciona una expresión de forma cerrada para P(y/do(x)) en términos de distribución de la probabilidad identificada:

Function: $\text{ClosedForm}(P(y \mid \hat{x}))$.

Input: Control query of the form $P(y \mid \hat{x})$.

Output: Either a closed-form expression for $P(y \mid \hat{x})$, in terms of observed variables only, or FAIL when the query is not identifiable.

1. If $(X \perp\!\!\!\perp Y)_{G_{\overline{X}}}$ then return $P(y)$.
2. Otherwise, if $(X \perp\!\!\!\perp Y)_{G_{\underline{X}}}$ then return $P(y \mid x)$.
3. Otherwise, let $B = \text{BlockingSet}(X, Y)$ and $Pb = \text{ClosedForm}(b \mid \hat{x})$; if $Pb \neq$ FAIL then return $\sum_b P(y \mid b, x) * Pb$.
4. Otherwise, let $Z_1 = \text{Children}(X) \cap (Y \cup \text{Ancestors}(Y))$, $Z_3 = \text{BlockingSet}(X, Z_1)$, $Z_4 = \text{BlockingSet}(Z_1, Y)$, and $Z_2 = Z_3 \cup Z_4$; if $Y \notin Z_1$ and $X \notin Z_2$ then return $\sum_{z_1, z_2} \sum_{x'} P(y \mid z_1, z_2, x') P(x' \mid z_2) P(z_1 \mid x, z_2) P(z_2)$.
5. Otherwise, return FAIL.

Conclusión

Los pasos 3 y 4 invocan la función BlockingSet(X, Y), que selecciona un conjunto de nodos Z que d-separa X de Y. Estos conjuntos se pueden encontrar en el tiempo polinomial.

El paso 3 contiene una llamada recursiva al algoritmo ClosedForm(b/do(x)), para obtener una expresión para el efecto causal P(b/do(x)).