AXIOMATIZING CAUSAL INFERENCE AND ITS EMERGING SIGNIFICANCE

Judea Pearl
UCLA

Cornell Symposium in honor of Joseph Halpern
June 2, 2023

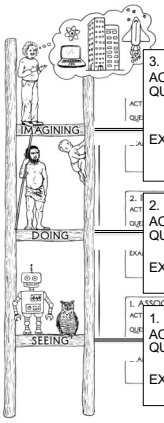
1

OUTLINE

1. The Ladder of Causation
2. Causal Models and Counterfactuals
3. Halpern Axiomatization of Counterfactuals
4. Why it matters
   - 4.1 Logical Equivalence of Structural and NR Framework
   - 4.2 Logical Equivalence ≠ Methodological Equivalence

2

3-LEVEL HIERARCHY



3. COUNTERFACTUALS
ACTIVITY: Imagining, Retrospection, Understanding
QUESTIONS: *What if I had done . . . ? Why?* (Was it X that caused Y? What if X had not occurred? What if I had acted differently?)
EXAMPLES: Was it the aspirin that stopped my headache? Would Kennedy be alive if Oswald had not killed him? What if I had not smoked the last 2 years?

2. INTERVENTION
ACTIVITY: Doing, Intervening
QUESTIONS: *What if I do . . . ? How?* (What would Y be if I do X?)
EXAMPLES: If I take aspirin, will my headache be cured? What if we ban cigarettes?

1. ASSOCIATION
ACTIVITY: Seeing, Observing
QUESTIONS: *What if I see . . . ?* (How would seeing X change my belief in Y?)
EXAMPLES: What does a symptom tell me about a disease? What does a survey tell us about the election results?

3

STRUCTURAL CAUSAL MODELS

Definition: A structural causal model is a 4-tuple $\langle V, U, F, P(u)\rangle$, where

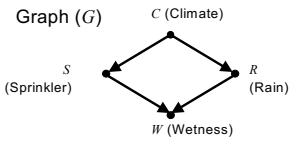
- $V = \{V_1,...,V_n\}$ are endogenous variables
- $U = \{U_1,...,U_m\}$ are background variables
- $F = \{f_1,...,f_n\}$ are functions determining V, $v_i = f_i(v, u)$ e.g., $y = \alpha + \beta x + u_Y$
- $P(u)$ is a distribution over U

$P(u)$ and F induce a distribution $P(v)$ over observable variables

4

A MODEL AND ITS GRAPH

Graph (G)



Model (M)

$$C = f_C(U_C)$$
$$S = f_S(C, U_S)$$
$$R = f_R(C, U_R)$$
$$W = f_W(S, R, U_W)$$

5

CAUSAL MODELS AND COUNTERFACTUALS

Definition:

The sentence: "Y would be y (in situation u), had X been x," denoted $Y_x(u) = y$, means:

The solution for Y in a mutilated model M_x, (i.e., the equations for X replaced by $X = x$) with input U=u, is equal to y.

The Fundamental Equation of Counterfactuals:

$$\boxed{Y_x(u) = Y_{M_x}(u)}$$

6

CAUSAL MODELS AND COUNTERFACTUALS

Definition:

The sentence: "Y would be y (in situation u), had X been x," denoted $Y_x(u) = y$, means:

The solution for Y in a mutilated model M_x, (i.e., the equations for X replaced by $X = x$) with input $U=u$, is equal to y.

- Joint probabilities of counterfactuals:

$$P(Y_x = y, Z_w = z) = \sum_{u:Y_x(u)=y,Z_w(u)=z} P(u)$$

7

CAUSAL MODELS AND COUNTERFACTUALS

Definition:

The sentence: "Y would be y (in situation u), had X been x," denoted $Y_x(u) = y$, means:

The solution for Y in a mutilated model M_x, (i.e., the equations for X replaced by $X = x$) with input $U=u$, is equal to y.

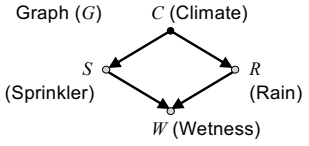
- In particular:

$$P(y|do(x)) \triangleq P(Y_x = y) = \sum_{u:Y_x(u)=y} P(u)$$

$$PN(Y_{x'} = y'|x, y) = \sum_{u:Y_{x'}(u)=y'} P(u|x, y)$$

8

DERIVING COUNTERFACTUALS FROM A MODEL



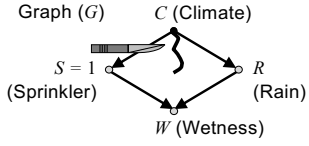
Model (M)

$$C = f_C(U_C)$$
$$S = f_S(C, U_S)$$
$$R = f_R(C, U_R)$$
$$W = f_W(S, R, U_W)$$

Would the pavement be wet HAD the sprinkler been ON?

9

DERIVING COUNTERFACTUALS FROM A MODEL



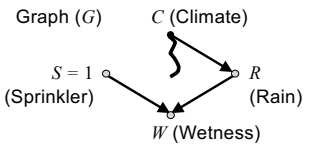
Mutilated Model ($M_{S=1}$)

$$C = f_C(U_C)$$
$$S = 1$$
$$R = f_R(C, U_R)$$
$$W = f_W(S, R, U_W)$$

Would the pavement be wet had the sprinkler been ON?
Find if $W = 1$ in $M_{S=1}$
Find if $f_W(S = 1, R, U_W) = 1$ or $W_{S=1} = 1$
What is the probability that we find the pavement is wet if we turn the sprinkler ON?
Find if $P(W_{S=1} = 1) = P(W = 1 \mid do(S = 1))$

10

DERIVING COUNTERFACTUALS FROM A MODEL



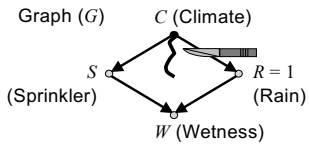
Mutilated Model ($M_{S=1}$)

$$C = f_C(U_C)$$
$$S = 1$$
$$R = f_R(C, U_R)$$
$$W = f_W(S, R, U_W)$$

Would it rain if we turn the sprinkler ON?
Not necessarily, because $R_{S=1} = R$

11

DERIVING COUNTERFACTUALS FROM A MODEL



Mutilated Model ($M_{R=1}$)

$$C = f_C(U_C)$$
$$S = f_S(C, U_S)$$
$$R = 1$$
$$W = f_W(S, R, U_W)$$

Would the pavement be wet had the rain been ON?
Find if $W = 1$ in $M_{R=1}$
Find if $f_W(S, R = 1, U_W) = 1$

EVERY COUNTERFACTAUL HAS A VALUE IN M

12

PROPERTIES OF STRUCTURAL COUNTERFACTUALS

Property 1 (Composition)
For any three sets of endogenous variables X, Y, and W in a causal model, we have

$$X_W(u) = x \Longrightarrow Y_{WX}(u) = Y_W(u).$$

13

PROPERTIES OF STRUCTURAL COUNTERFACTUALS

Property 1 (Composition)
For any three sets of endogenous variables X, Y, and W in a causal model, we have

$$X_W(u) = x \Longrightarrow Y_{WX}(u) = Y_W(u).$$

Property 2 (Effectiveness)
For all sets of variables X and W, $X_{XW}(u) = x$.

14

PROPERTIES OF STRUCTURAL COUNTERFACTUALS

Property 1 (Composition)
For any three sets of endogenous variables X, Y, and W in a causal model, we have

$$X_W(u) = x \Longrightarrow Y_{WX}(u) = Y_W(u).$$

Property 2 (Effectiveness)
For all sets of variables X and W, $X_{XW}(u) = x$.

Property 3 (Reversibility)
For any two variables Y and W and any set of variables X,

$$(Y_{XW}(u) = y)\ \&\ (W_{Xy}(u) = w) \Longrightarrow (Y_X(u) = y).$$

15

HALPERN'S COMPLETENESS THEOREMS

Theorem 7.3.5 (Recursive Completeness)
Composition, effectiveness, and recursiveness are complete.

Theorem 7.3.6 (Completeness)
Composition, effectiveness, and reversibility are complete for all causal models.

16

TWO DIALECTS FOR CAUSAL INFERENCE

Observed: $P(X, Y, Z,...)$
Conclusions needed: $P(Y_x=y)$, $P(X_y=x \mid Z=z)$...

How do we connect observables, X, Y, Z,…
to counterfactuals Y_x, X_z, Z_y,… ?

N-R model	Structural model
Counterfactuals are primitives, new variables	Counterfactuals are derived quantities
Super-distribution $P^*(X, Y, \ldots, Y_x, X_z, \ldots)$	Subscripts modify the model and distribution $P(Y_x = y) = P_{M_x}(Y = y)$
Assumptions: Composition (SUTVA)	

17

INFERENCE IN POTENTIAL-OUTCOME FRAMEWORK

Define:	Express the target quantity Q as a counterfactual formula
Assume:	Formulate causal assumptions using composition and the distribution: $P(X\|Y, Z, Y(1), Y(0))$
Identify:	Determine (algebraically) if Q is identifiable.
Estimate:	Estimate Q if it is identifiable; approximate it, if it is not.

18

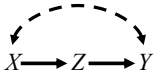
FORMULATING ASSUMPTIONS THREE LANGUAGES

1. English: Smoking (X), Cancer (Y), Tar (Z), Genotypes (U)

2. Counterfactuals: $Z_x(u) = Z_{yx}(u)$
$X_y(u) = X_{zy}(u) = X_z = X(u),$
$Y_z(u) = Y_{zx}(u),$
$Z_x(u) \perp\!\!\!\perp \{Y_z, X\}$

Not too friendly:
Consistent?, complete?, redundant?, arguable?

4. Structural:



19

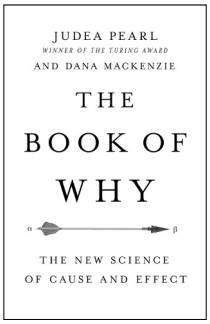
CONCLUSIONS

1. The SCM and PO frameworks are logically equivalent
   1.1 A Theorem in one is a theorem in the other
   1.2 Starting with same assumptions, a SCM researcher will never disagree with a PO enthusiast

2. Logical Equivalence ≠ Methodological Equivalence
   2.1 Friendly models can be as expressive and as safe as unfriendly ones

20

For a trailer, click WHY on my home page.
For discussions: @yudapearl



21