TECHNICAL REPORT
R-470
April 2017

The Eight Pillars of Causal Wisdom (Lecture notes for the UCLA WCE conference, April 24, 2017)

Judea Pearl
University of California, Los Angeles
Computer Science Department
Los Angeles, CA, 90095-1596, USA
(310) 825-3243
judea@cs.ucla.edu

Abstract

This title was chosen partly to echo Stigler's "The Seven Pillars of Statistical Wisdom" and partly to counter Angrist and Pischke's advocacy of model-blind economics. To Stigler it says: Statistics without Causality is at least one pillar short on wisdom. To model-blind economists it says: You can't do science without a microscope.

As we enter an era of machine learning and "Big Data," it is important to realize that we are also entering an era of a Causal Revolution, which now casts a sobering light on the theoretical limitations of data-centric methodologies, and, hence, on the future direction of data science.

1 Introduction

How effective is a given treatment in preventing a disease? Was it the new tax break that caused our sales to go up? What is the annual health-care costs attributed to Obesity? Can hiring records prove an employer guilty of sex discrimination? I am about to quit my job, but should I?

The peculiar nature of these questions is that they are constantly being asked, both by researchers, policy makers and ordinary folks, yet, until very recently they could not have been articulated, let alone answered, using the available scientific methods. These questions all concern cause and effect relationships which, unlike geometry, mechanics, optics or probabilities, have been denied the benefits of mathematical analysis until very recently.

To appreciate the extent of this denial, readers would be stunned to know that only a few decades ago scientists were unable to write down a mathematical equation for the obvious fact that "mud does not cause rain." Even today, only the top echelon of the scientific community can write such an equation and formally distinguish "mud causes rain" from " rain causes mud." And you would probably be even more surprised to discover that your favorite college professor is not among them. Try it!

Things have changed dramatically in the past three decades, A powerful and transparent mathematical language has been developed for managing causes and effects, accompanied by a set of tools that turn causal analysis into a mathematical game, not unlike solving algebraic equations, or finding proofs in high-school geometry. These tools permit us to express causal questions formally, codify our existing knowledge in both diagrammatic and algebraeic forms, and then leverage our data to estimate the answers. Moreover, the theory warns us when the state of existing knowledge or the available data are insufficient to answer our questions; and suggests additional sources of knowledge or data to make the questions answerable.

I call this transformation "The Causal Revolution," and the mathematical framework that led to it I call "Structural Causal Models (SCM)."

The SCM invokes three mathematical objects:

1. Graphical models,

2. Structural equations, and

3. Counterfactual and interventional logic.

Graphical models serve as a language for representing what we know about the world, counterfactuals articulate what we want to know, while structural equations serve to tie the two together in a solid semantics.

In this talk, I will provide a bird's eye view of the eight most significant accomplishments of the SCM framework and discuss the unique contribution that each pillar brings to the research community.

2 The Eight Pillars

2.1 Graphical models for prediction and diagnosis

I will describe how Bayesian Networks emerged in the late 1980s and how graphical models became the standard method of handling uncertainty in expert systems and machine learning. In particular, I will discuss how graphs and probabilities, two total strangers, came together to create a reasoning machine for revising beliefs in light of new evidence.

2.2 The control of confounding

I will describe how confounding, the major obstacle to drawing causal inference from data, had been demystified and totally "deconfounded" using a graphical criterion called "back-door." In particular, how the task of selecting an appropriate set of covariates for control of confounding has been reduced to a simple "roadblocks" puzzle manageable by rank and file investigators.

2.3 *Do*-calculus – An all-seeing oracle for predicting the effects of policies and interventions

I will describe the workings of a symbolic engine that determines, for any given model, whether a randomized experiment can be replaced by an observational study and, if so, how. Using this engine, called *do-calculus*, the effect of policy interventions can be predicted by a simple mathematical exercise, resembling theorem proving in high-school geometry. This oracle can be applied to time-varying policies and time varying observations, including problems for which control of confounding is infeasible.

2.4 The algorithmization of counterfactuals

Counterfactual analysis deals with behavior of specific individuals, identified by a distinct set of characteristics. For example, given that Joe's salary is $Y = y$, and that he attended $X = x$ years of college, what would Joe's salary be had he had one more year of education.

One of the crown achievements of SCM was to formulate counterfactual reasoning within the same graphical representation that researchers use to encode scientific knowledge. Every structural equation model determines the truth value of every counterfactual sentence. Therefore, we can determine analytically if the probability that such a sentence is true is estimable from experimental or observational studies, or combination thereof.

Of special interest to researchers are counterfactual questions concerning "causes of effects," as opposed to "effects of causes." For example, how likely it is that Joe's swimming exercise was a necessary (or sufficient) cause of Joe's death.

I will explain why the graphical nature of the models ensures the transparency of modeling assumptions, how ordinary investigators can now judge the plausibility of those assumptions, identify their testable implications, and submit them to the scrutiny of data. I will also demonstrate how one can avoid the opacity of "ignorability" assumptions and replace them with meaningful substantive knowledge.

2.5 Mediation analysis and the assessment of direct and indirect effects

Mediation analysis concerns the mechanisms that transmit changes from a cause to its effects. I will explain why the detection of intermediate mechanism is essential in many research problems and why counterfactual logic must be invoked to facilitate this detection. In particular, I will demonstrate how the graphical representation of counterfactuals enables us to define direct and indirect effects and to decide when these effects are estimable from data, or experiments.

2.6 External validity and sample selection bias

The validity of every experimental study is challenged by disparities between the experimental and implementational setups. Since participation in randomized trials cannot be mandated, we cannot guarantee that the study population would be the same as the population intended for treatment. Fortunately, the *do*-calculus discussed above now offers a complete methodology for overcoming this source of bias. This non-parametric methodology goes way beyond the standard methods of meta-analysis and recalibration. I will describe briefly the tools available for establishing external validity and for controlling bias due to non-representative samples. (Elias Bareinboim will elaborate on these issues in his lecture).

2.7 Missing data

Problems of missing data plague every branch of experimental science. Respondents do not answer every item on a questionnaire, sensors fade as environmental conditions change, and patients often drop from a clinical study for unknown reasons. The rich literature on this problem is almost totally wedded to the statistical, model-blind paradigm of Rubin and Little. Recently, the problem has been addressed from a totally new perspective, using causal diagrams to model the process of data missingness. With the help of these models we can now understand (and formalize) the conditions under which causal and probabilistic relationships can be recovered from incomplete data and, whenever the conditions are satisfied, we can construct an unbiased estimate of the desired relationship. (Karthika Mohan will give a full lecture on these exciting results.)

2.8 Causal discovery

The ability to detect and enumerate the testable implications of a given graphical model, opens the possibility of inferring, with very mild assumptions, the set of models that are compatible with the data. I will summarize the basic premises behind this method of discovery, its promises and its weaknesses.

Summary

The SCM framework stole causal inference from Mt. Olympus and brought it down for ordinary researchers to use.

References

BAREINBOIM, E. and PEARL, J. (2016). Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences* **113** 7345–7352.

BAREINBOIM, E., TIAN, J. and PEARL, J. (2014). Recovering from selection bias in causal and statistical inference. In *Proceedings of the Twenty-eighth AAAI Conference on Artificial Intelligence* (C. E. Brodley and P. Stone, eds.). AAAI Press, Palo Alto, CA, 2410–2416. Best Paper Award.

Pearl, J. (2014a). The deductive approach to causal inference. *Journal of Causal Inference* **2** 115–129.

Pearl, J. (2014b). Interpretation and identification of causal mediation. *Psychological Methods* **19** 459–481.

Pearl, J. (2015). Trygve Haavelmo and the emergence of causal calculus. *Econometric Theory* **31** 152–179. Special issue on Haavelmo Centennial.

Pearl, J. and Bareinboim, E. (2014). External validity: From *do*-calculus to transportability across populations. *Statistical Science* **29** 579–595.

Pearl, J. and Mohan, K. (2014). Recoverability and testability of missing data: Introduction and summary of results. Tech. Rep. R-417, <http://ftp.cs.ucla.edu/pub/stat_ser/r417.pdf>, Department of Computer Science, University of California, Los Angeles, CA.

(Electronic versions of these and other papers can be viewed here:
⟨http://bayes.cs.ucla.edu/csl_papers.html⟩.)