On Tim Maudlin's [Review of *The Book of Why*]

This page links to a review of *The Book of Why* by Tim Maudlin, published September 4, 2019, at the BostonReview.net. Over all, the review is quite supportive of the main ideas developed in the book, and it exposes those ideas in a lucid, accessible way, enriched by the extensive scientific background of the reviewer. It has also generated a lively discussion on my Twitter page, which I would like to summarize here and use this opportunity to clarify some not-so-obvious points in the book, especially the difference between Rung Two and Rung Three in the Ladder of Causation

There are two main points to be made on the relationships between the two rungs: interventions and counterfactuals.

(1) Although you CAN phrase actions and interventions (rung two) in terms of counterfactuals, the important and surprising fact is that you do not have to. As long as we are talking about population averages rather than individuals, we can do it with simple diagrammatic procedures (erasing arrows). This is demonstrated vividly in Causal Bayesian Networks (CBN) which enable us to compute theaverage causal effects of all possible actions, including compound actions and actions conditioned on observed covariates, while invoking no counterfactuals whatsoever. Each CBN lends itself to empirical verification using randomized trials, free from of the "multiple worlds" criticism that empiricists often cite against counterfactuals. For definitions and further details see (Pearl 2000 (Ch. 3), Dawid 2000, Pearl et al. 2015 (Primer)).

The fact that we CAN state an intervention in terms of a counterfactual (e.g., Rubin 1974, Robins 1986, Pearl 1995) has given some people the impression that this is the ONLY way to define an intervention. Hence there is no difference between counterfactuals and "causality," or rung two and rung three. Indeed, the distinction is absent in the books of Imbens and Rubin (2015), Hernán and Robins (2019) and, evidently, the philosophical literature that informed Tim Maudlin.

(2) On the converse side of the coin, there ARE counterfactual statements that we cannot phrase in the language of the do-operator. We make this very clear in the book when we talk about the struggle to express mediation in this way. The practical ramification of this mathematical fact is that certain counterfactual questions cannot be answered using experimental studies, among them questions about mediation (e.g., is an employer guilty of sex discrimination in hiring] and causes of effect [e.g., Was drug D the cause of Joe's death] The profound difference between experiments and counterfactuals are further demonstrated in https://ucla.in/2Qb1h6v and https://ucla.in/2G2rWBv.

A second reservation I have about Maudlin's review concerns his opinion that "Pearl could have saved himself literally years of effort had he been apprised of this work" [philosophy teaching at Pittsburgh University in the 1980's]. I have challenged my 22,000 twitter followers to name one idea from pre-1990 philosophy that I have missed, and that could have saved me a day or two. I believe Maudlin was unaware of the fact that, in the 1980's, I was intimately familiar with the Tetrad project of Clark Glymour, as well as with much of the philosophical literature on causality, from Reichenbach and Salmon, to Lewis, Skyrms, Suppes and Cartwright.

I hope this review entices philosophy-minded readers to take a look at how causality has been transformed since the writings of these philosophers. In particular, how counterfactuals have been algorithmatized and how our understanding of the limits of empirical science is illuminated through the Ladder of Causation.

Judea Pearl