Discussion of Causal inference by using invariant prediction: identification and confidence intervals, by Peters, Buhlmann and Meinshausen

The authors have produced a stimulating paper, which will be of interest not only to statisticians, but also to people working in other communities, such as artificial intelligence. The authors note that if one can identify all the direct causes / causal predictors of a response variable then the distribution of this variable conditioned on these predictors will be invariant under manipulation of other variables in the system. This could be thought of as a direct consequence of the directed local Markov property that a variable is independent of its non-descendants given its parents (see for example [Lauritzen, 2001]). They then look for such invariance across different environments in order to identify these predictors.

The authors have shown that the set of causal predictors is identifiable when manipulations of the system are of certain types (Theorem 2), including the rudimentary *do* interventions of Pearl [Pearl, 2000]. However, they also make the assumption (in for example section 7.1) that the exact nature of the interventions is unknown. If this is indeed the case, how probable is it that the interventions are of these types? An urgent task is to demonstrate that the set of predictors is identifiable for a much wider class of interventions – if those listed turn out to be the only ones that allow this set to be identified, then the work in this paper, however interesting, may turn out to be of limited use. I would like to propose investigating the following types of intervention as being among those of interest:

- Manipulating collections of variables to specific values, where there is not at least one single *do* intervention on each non-response variable.
- *Stochastic* manipulations which assign a new probability distribution over the outcomes of manipulated variables.
- *Functional* manipulations $Do\ X = g(W)$ for some set of variables W.

We could of course also consider what might be termed *stochastic functional* manipulations.

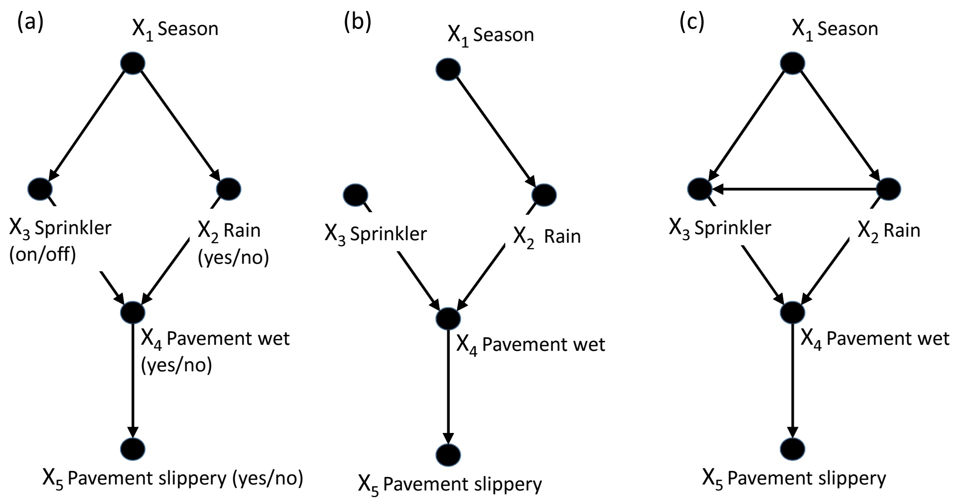

Figure 1: Example of a *functional* manipulation

I will concentrate here on functional manipulations. So consider the *Sprinkler* example from [Pearl, 2000], a DAG for which is given in Figure 1 (a). Here, using the adapted methodology of section 6.1, we have SEMs: $X_1 = f(\varepsilon_1)$, $X_2 = f_2(X_1, \varepsilon_2)$, $X_3 = f_3(X_1, \varepsilon_3)$,

$X_4 = f_4(X_2, X_3, \varepsilon_4)$, $X_5 = f_5(X_4, \varepsilon_5)$. The *do* intervention *Put sprinkler on* removes the edge $X_1 \to X_3$ (as in Figure 1 (b)), and hence X_3 is no longer a function of X_1. But we could consider a manipulation such as *If it is Summer put the sprinkler on; if it is not Summer and it is raining put the sprinkler off* [Thwaites, 2013]. Here, instead of removing the edge $X_1 \to X_3$, we need to add an edge $X_2 \to X_3$ as in Figure 1 (c), since whether the sprinkler is on depends on both the season and whether it is raining. So a possible SEM for this is $X_3 = f'_3(X_1, X_2)$, implying a deterministic relationship between X_1, X_2 and X_3. But what happens to the sprinkler if it is not Summer and not raining?

It is not immediately apparent whether these kind of scenarios will always satisfy the assumptions stated in the paper, and if they do, whether the set of causal predictors will always be identifiable. In this particular example this might not be an issue since the parents of the probable response variables X_4 and X_5 remain unchanged.

The authors have extended their ideas to the non-linear case. The *Sprinkler* example here which uses discrete variables, suggests to me the further extension to cases where the methodology must necessarily be non-parametric. I would also like to draw attention to the (still relatively small) collection of books and papers on causality which argue that *causes* are more naturally thought of as *events*, rather than random variables (see for example [Shafer, 1996, Dawid, 2000, Thwaites et al., 2010]). Is the analysis in this paper compatible with this interpretation?

As befits a Discussion paper, this article provides plenty of opportunity for debate, argument and further research. It is therefore with great pleasure that I propose a vote of thanks to the authors.

References

[Dawid, 2000] Dawid, A. P. (2000). Causal inference without counterfactuals. *Journal of the American Statistical Association*, 95:407–448.

[Lauritzen, 2001] Lauritzen, S. L. (2001). Causal inference from graphical models. In Barndorff-Nielsen, O. E. et al., editors, *Complex Stochastic Systems*. Chapman and Hall.

[Pearl, 2000] Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge.

[Shafer, 1996] Shafer, G. (1996). *The Art of Causal Conjecture*. MIT Press.

[Thwaites, 2013] Thwaites, P. A. (2013). Causal identifiability via Chain Event Graphs. *Artificial Intelligence*, 195:291–315.

[Thwaites et al., 2010] Thwaites, P. A., Smith, J. Q., and Riccomagno, E. M. (2010). Causal analysis with Chain Event Graphs. *Artificial Intelligence*, 174:889–909.