Causal Cognition Midterm Essay

(2) What are the key differences between inferences at different rungs of Pearl's ladder of causation? Explain with reference to examples.

Nathan A. Sharp

Edinburgh University

Causal Cognition Midterm Essay

(2) What are the key differences between inferences at different rungs of Pearl's ladder of causation? Explain with reference to examples.

Pearl's ladder of causation is a simple and powerful theoretical taxonomy differentiating progressively more complex classes of causal reasoning in intelligent agents. This essay will unpack the details of Pearl's ladder and investigate its reasoning and relevance.

Each rung of Pearls ladder represents a structure of causal reasoning as follows, (1) Association -> (2) Intervention -> (3) Counterfactual, with increasing model complexity as you ascend the ladder. Pearl claims his ladder is, (i) mathematically rigorous in the claim that higher rung reasoning fundamentally cannot be reduced to the symbolic language of lower rungs as it requires novel representation, (ii) combined, are the default framework humans (and animals) actually use reason about the world (as opposed to probabilistic models), and (iii) that its higher rung reasoning is a hallmark of human intelligence and a major factor separating humans from animals and present artificial intelligence. This essay will also broadly touch on these claims when possible.

The bottom rung of the ladder, Association "entails detection of regularities in our environment" (Pearl, 2018). It is fundamentally observational as you are not intervening with the world beyond processing the data to infer, update and restructure possible causal dependencies. For example you see someone coughing in public and infer this could be a symptom of Covid-19. If you then then observed them smoking a cigarette you might update your causal model to reduce the likelihood that their cough was caused by Covid.

The second rung of the ladder, Intervention involves causal reasoning about acting in the world. What if I wear a mask in public (X)? how will it effect my chance of catching Covid (Y)? Pearl encapsulates this class of inference and how it differentiates from purely observational inferences in what he calls 'do' notation which satisfies

$$P(Y|X) \not\equiv P(Y|\, do(X)).$$

This reads the probability of observing Y given the observation of X in the wild ($P(Y|X)$) is not necessarily equivalent to the probability of observing Y as the result of a (randomised) intervention ($P(Y|\,do(X))$. With respect to our example, the correlation between people freely wearing a mask and catching Covid ($P(Covid|\,Mask)$ will likely be different to the correlation if a wearing a mask is mandated ($P(Covid|\,do(Mask))$). Consider with respect to our example that a generally more cautious person might be more likely to wear a mask and so be more likely not to catch Covid for reasons correlated with but not caused by their mask wearing.

This inference class of causal intervention and its do-operator is foundational to Pearl's worldview– the primacy and 'mathematisation' of causal ontology. Its one of the hallmarks of causal reasoning, separating it from its merely statistical cousin, probabilistic reasoning. Pearl claims this ability to incorporate the effects of these 'do' interventions as a change to the causal structure rather than a 'dumb' incorporation mixed into a statistical probability vector is what makes causal models more stable to change than probabilistic ones "(causal models) allow us to make predictions under a wider range of circumstances, including circumstances in which things are taken appart, reconfigured, or undergo spontaneous change." (Pearl, 2009). Pearl uses this 'do more with less' argument in part to support his thesis that causal models (as opposed to probabilistic ones) are the de facto currency of human reasoning.

The third and highest rung of Pearl's ladder is counterfactual causal reasoning. It entails imaginative 'what if?' reasoning about things that might have happened but didn't or did happen but might not have. For example thinking about how not going to a party might have prevented you from catching covid requires the ability to think about a world that doesn't exist and never will. Pearl asserts of counterfactual reasoning that "(it) is the building blocks of moral behavior as well as scientific thought. The ability to reflect on one's past actions and envision alternative scenarios is the basis of free will and social responsibility" (Pearl, 2018).

Pearl thinks that the existence and sophistication of human counterfactual reasoning is one of the major factors separating human intelligence from animal and

present-day artificial intelligence. In the case of artificial intelligence I think his case was very strong at the time of his original breakthrough on causality (around the turn of the century) and even the subsequent generation of powerful deep neural networks were very much rung-1 'dumb' yet large statistical inference machines. However recent advances areas such as reinforcement learning seem to me much more in line with Pearls description of reasoning at higher rungs as these algorithms embody agents that have to make decisions on acting in the world (intervention) under the consideration their previous decisions and generate novel strategies (counterfactual). I also suspect animals are more capable of reasoning on all three rungs than we imagine, showing capabilities of planning, empathy, etc. Even if we accept claim (i) of the fundamental irreducibly of each rung I'm more inclined to the view that many creatures aka not the defining factor in separating human intelligence

In conclusion, Pearl's ladder is an interesting concept powerful tool in his wider causality toolbox and certainly central to his mission statement of "explaining how to answer causal queries and what information is needed to answer them." (Pearl, 2018) Personally I don't think its the most important part of his worldview but its simplicity coupled with relevance to a wide range of phenomena makes it a more than worthwhile model for any psychologist to keep in mind.

References

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd). Cambridge University Press.

Pearl, J. (2018). *The book of why*. Basic Books.