Comments on Ned Hall's
"Causation and the Aims of Inquiry"

Eddy Keming Chen
eddy.chen@rutgers.edu

February 18, 2016

I'd like to begin by thanking Professor Hall, for his interesting and provocative paper on causation and the aims of scientific inquiry. In the following, I shall provide three groups of comments, which are based partly on his paper and partly on his talk.

1 General Theme: Metaphysical Foundations for Causal Modeling

The general theme of Hall's paper is to provide the appropriate metaphysical foundations for the interventionist account(s) of causation. Interventionism is an increasingly popular approach in the social and medical sciences. It has many virtues. For example, it provides diagrams and mathematical models that can help the scientists separate causal relations from mere statistical correlations. But as Hall suggests (rightly), the contemporary literature on interventionism, causal modeling, structural equations does not address their metaphysical foundations.

That seems correct. But we might wonder *why* interventionism needs such metaphysical foundations. The background assumption, in much of the literature, and in Hall's paper, is that there is no fundamental obtaining causal relation in the actual world. Indeed, the interventionist account itself seems to suggest, according to Pearl, that causality disappears when we want to model the entire universe as a whole system, which is what fundamental metaphysics is about and what fundamental physics aspires to describe. Given that, fundamentally there is no causation, what do we ask for when we seek the metaphysical foundations of interventionism?

Hall's answer, at least partly, is this: causal models often come with certain graphs and some set of structural equations, describing the dependencies among the causal variables. However, in a given concrete situation, what makes one set of structural equations the correct one? Typically, people give the answer in terms of interventions, which are themselves explained in terms of further causal models! Hence, the task of finding metaphysical foundations of causal modeling involves finding non-circular explanations for the their correctness conditions.

But we can ask a further question. Causal talks are useful and convenient ways of describing many things around us: medium sized dry goods behave according to our folk physics around sentient animals such as ourselves. They can be *extrapolated*

to circumstances more removed from our immediate experiences. But why suppose such extrapolation, and the subsequent application of the folk notion of causation, have correctness conditions? And if they do, why would there be unique correctness conditions? After all, Hall's aim of a successful account of causation is utility, and not truth (at least not primarily truth). It could well be that causal models are correct as long as they serve some pragmatic purposes of making things happen.

2   Hall's Proposal: The Theory of Localized Dependence Structure

In Section 2 of the paper, Hall provides a positive proposal—the Theory of Localized Dependence Structure. On this theory: "the world possesses a localized dependence structure, constituted by the totality of facts about how conditions at different spatially and temporally bounded places and times depend on conditions at other such places and times," and the causal structure is the localized dependence structure. (My reading is that the regions also have to be connected.) Hall cashes out "conditions" and "dependence" in the following way. For conditions, we need to specify, for a given inquiry, the scale, the aspects, the level of precision, and the similarity metric of that inquiry. Such freedom of choice explains why we have different branches of sciences, operating on different domains and levels of reality. For the dependence relation, we understand it in terms of counterfactual conditionals. Note: Hall is not trying to capture their meanings in everyday discourse; he is explaining how we *should* understand them, given their *role* in specifying the notion of localized dependence structure.

The counterfactuals described in the following schema

Schema 1 Had conditions C1 obtained in region R1, C2 would have obtained in region R2;

are evaluated by a Maudlin-style semantics. In the easiest case, when R1 only spans one instant in time (or occupying at most one space-like hypersurface), and when C1 specifies a precise fundamental physical state, assuming that determinism is true, that R2 is to the future of R1 while not assuming that R2 is instantaneous, we proceed in 3 steps: (1) consider a possible complete physical state of our world at t (simultaneous with R1 wrt to the space-like hypersurface) that is exactly like its actual state except that C1 obtains in R1, (2) apply the physical laws of our world to that state, (3) check the final unique state after the dynamical evolution—the conditional is true if C2 obtains in R2; the conditional is false if C2 does not obtain in R2. That's how we should evaluate the counterfactuals in the easiest case.

We can think of the theory as a version of the interventionist account of causation, one in which intervention on a system is crucial for evaluating the obtaining of causal relations that are spelled out in terms of counterfactual conditionals connecting what happens here to what happens there. These conditionals, on Hall's view, capture the outcomes of perfectly controlled experiments. Just as an experimenter, you intervene by wiggling localized conditions in some regions of the world and observe what happens in other localized regions. The causal structure is just the localized dependence structure. If so, then this account is highly useful, because it clarifies the content of what he calls

Inquiry's Aim: Scientific inquiry aims to discover and describe the causal structure of the world.

That is:

Inquiry's Aim+TLDS: Scientific inquiry aims to discover and describe the localized dependence structure of the world.

Here we see a bright virtue of Hall's theory, as it clarifies the content of the Inquiry's Aim.

I'd like to raise two questions about this proposal. First: an immediate question about the identification between these two structures: the localized dependence structure and the causal structure that scientific inquiry aims to discover and describe. Does it imply, that a large chunk of contemporary science is not serious scientific inquiry? If not, why not? There seem to be many parts of sciences in which the inquiry is not about dependence, or about localized dependence. Let us take fundamental physics as an example. We have learnt from Bell's famous 1964 theorem and the subsequent empirical violations of Bell Inequalities that our world is non-local. That is, there are fundamental physical dependence structure (such as the entanglement relation) that is non-local. More explicitly, our quantum theory, understood either in deterministic de-Broglie-Bohm formulation or in stochastic GRW formulation, is dramatically non-local. Such non-locality is evident in the wave function of the universe, which is "responsible" for the distribution of matters in space-time. Quantum theory is non-local. Our fundamental physical theory aims to capture the predictions of quantum theory. So it is also non-local. A large chunk of scientific inquiry has been aiming at writing down various wave functions of entangled physical systems and ultimately the wave function of the universe (quantum cosmology). Since these are not aimed at describing localized dependence structure, Hall's proposal seems to disqualify them as serious scientific inquiry. I'm sure Hall's understanding of the thesis is more nuanced; so I invite him to say more during the discussion.

The second question is about the asymmetry of dependence: future depends on the past but not much vice versa; we can wiggle things now to make changes to the future but not much to the past. We know that *that* is typically true of our part of the universe (and it probably deserves celebration by embodied agents such as ourselves). The question is why. But Hall's theory seems to build it in by hand, as his semantics only apply to counterfactuals with R2 to the future of R1. The asymmetry of dependence is a puzzling feature, given that there is nothing that stands between us and the past, and that all the fundamental laws of physics that have been seriously considered (Newtonian, Maxwellian, and Schrodinger dynamics) are time-reversal symmetric (modulo the controversy surrounding David Albert's *Time and Chance* Chapter 1). It would be more desirable, for example, if the asymmetry emerges from something more natural. Here we a have choice point: (A) incorporate the Past Hypothesis a la Boltzmann-Feynmann-Penrose-Albert-Loewer into part of the laws, (B) incorporate Maudlin's primitive arrow of time (and explain how it gives rise to the asymmetry of dependence), and (C) develop new dynamical equations with built-in asymmetry. In any case, there might be interesting work to be done regarding how the choice about the fundamental asymmetry fits in to the general picture of providing

metaphysical foundations for causal modeling. That would presumably give rise to different metaphysical foundations for interventionism.

3 Protracted Causes: A Celebration of Freedom?

One worry Hall raises for his own account is about causes that are extended in time, or "protracted causes." Recall that although physics does not impose much synchronic constraints on the distribution of matter, it imposes much diachronic constraints on that. (That is, setting aside synchronic constraints such as the Pauli Exclusion Principle, the divergence of the magnetic field equaling zero, quantum equilibrium, and constraints in the presence of CTCs; actually quite a long list.) One problem is that brute interventions in the temporally extended R1 would sometimes violate fundamental physical laws. If we intervene at the first instant of R1, we can't guarantee that it instantiates C1 for the interval.

Recall the case (IF TIME PERMITS): Suzy breaks the window at 11:45, Billy comes to the scene with window-breaking intentions at 12:15 only to be disappointed at sight of the broken window. We'd like to know: what localized conditions of the world depend on the window's being broken during 12:00-12:30. We can't just intervene and make the window unbroken at 12:00, for that's insufficient to guarantee its being unbroken during 12:00-12:30. For that, we have to change Billy's intentions. Suppose we also change his intentions at noon. But Billy is not part of the localized region R1 yet! He's only coming in at 12:15! *So this suggests that we can't confine our intervention to conditions within R1. But since this opens up many ways of intervening outside R1, Hall thinks that we need to find a principled way (not relying on whatever is psychological salient) for the theory to be plausible.* A similar lesson applies to those causes that are monitored by their environments, such as the lunch conversation case.

That brings me to the final question: Why do we need a general, universal, and principled way to intervene on protracted causes? It seems like by his own lights, Hall's theory needs no such requirements. Instead, he can celebrate the freedom of choice. Recall that the theory of localized dependence structure already comes equipped with certain freedom of choice: it allows the conditions to be specified according to the scale, the aspects, the level of precision, and the similarity metric needed for the particular inquiry's aim. Wouldn't a carefully designed inquiry be aimed at something particular such that when we specify the right conditions, we would also make the choices for what is the most salient way to intervene outside the localized region? For example, we can further require the conditions to be specified according to the aim of inquiry and according to a set of variables, namely, the variables that we are interested in manipulating. This seems much more consistent with the spirit of Hall's Theory of Localized Dependence Structure.

I am sure that Ned has many thoughts about these questions. It was very rewarding for me to work through Ned's paper and discuss with him before his talk. I hope that these comments can help the audience to dig deeper into your ideas in discussion, which we shall turn to now (or after the break). Thank you.