DISSERTATION ABSTRACT
Tyrus Fisher

My dissertation, *Counterfactual Dependence and Causation's Arrows*, primarily concerns issues to do with Modus Ponens and its relations to other inference principles involving conditionals. The dissertation delivers three important results: (1) a semantics that handles both backtracking and counterlegal counterfactuals; (2) A defense of Modus Ponens for counterfactuals and a diagnosis of why apparent counterexamples are attractive; and (3) a normative constraint on semantic theories, which brings out the invalidity of an almost universally accepted inference pattern—the inference from the indicative to the material conditional. I elaborate below.

I begin by highlighting an overlooked limitation of previously-available semantic theories of counterfactuals—those that handle counterlegals cannot handle backtrackers and vice versa. The difficulty here is principled: Roughly, to handle backtrackers a semantics must hold laws (or some general principles) fixed, but to handle counterlegals, some allowance for law-breaking must be made. This work is published in Fisher (2016b). My first result is that both backtracking and counterlegal counterfactuals can be interpreted within a single semantic framework, which I provide. To elaborate, a semantics for counterfactuals should be able to appropriately evaluate conditionals such as:

(B) If Scotty were in Canberra now, he wouldn't have joined us for dinner this evening (here in California).

and

(C) Even if the temperature hadn't risen, that bottle containing that beer would have burst had its internal pressure risen above 150 psi.

(B) is a *backtracking counterfactual*—it tells us how an earlier time would be different had a later time been different. (C) is a *counterlegal*—its antecedent is inconsistent with some set of laws. Oddly, for other theories if they can handle backtrackers, they cannot handle counterlegals.

The surprising key to the solution amounts to thinking of a "normal" (i.e. non-counterlegal) counterfactual as a special kind of counterlegal rather than thinking of the latter kind as a special case of the former. The semantics I develop interprets a given counterlegal relative to a causal-model, M, representing the causal system the counterlegal is about, and relative to a class of models that are *minimally-illegal* (relative to M). Loosely, a causal-model is minimally-illegal just in case it represents a minimal violation of M's laws while satisfying the antecedent of interest and maximally preserving M's assignment of values to its variables. Non-counterlegals are then treated as degenerate cases of counterlegals—cases in which the set of violated laws is empty. In such degenerate cases, my semantics agrees with Hiddleston's (2005), a semantics that does well with backtrackers. This work is published in Fisher (2016b).

The next two results bear on the validity of Modus Ponens. I provide counterexamples to influential semantic theories that invalidate Modus Ponens when applied to right embedded counterfactuals—the interventionist theories of Pearl 2000 and Briggs 2012—and use these counterexamples to pinpoint what goes wrong with the intuitive apparent counterexamples to Modus Ponens used to buoy the aforementioned formal accounts. Consider, for example:

(M) If the match had lit, then if it hadn't been struck, it would have lit.

There are natural contexts relative to which this conditional comes out intuitively false. But interventionist theories admit no falsifying model.

I go on to show that the reason interventionist theories cannot adequately handle such conditionals is exactly the reason they cannot handle backtrackers, and, moreover, the reason they invalidate Modus Ponens. My counterexamples thus provide a route to a philosophical response to the interventionist's proposed counterexamples to Modus Ponens. They also show that the property of interventionist theories that makes them unable to interpret backtrackers and that invalidates Modus Ponens also makes them unable to adequately interpret some non-backtrackers. These points are made in Fisher (2016a).

In the final part of my dissertation, I argue that an apparently valid inference that has guided much theorizing about indicative conditionals is, in fact, invalid, This is the inference from the indicative conditional to its material counterpart. I show that the apparent validity of this inference is only supported by illicitly encoding metaphysically-prejudicial assumptions in one's semantic framework. These are assumptions that rule out certain extra-semantic theses such as indeterminism. But whether indeterminism holds or not is an extra-semantic matter, and there are forceful arguments from indeterminism (plus certain other coherent extra-semantic theses) to the failure of bivalence for future contingents. I then utilize the premise that a semantics should not rule out coherent extra-semantic theses and, so, admit models that give rise to truth-value gaps and argue for the conclusion that the indicative conditional does not entail its material counterpart. Here is a case to support this failure of entailment:

Consider a world with some genuine chanciness, and consider a coin flipping device wired so that a coin will be flipped just in case a sufficient quantity of radioactive material decays so that there is some objective chance that the coin will be flipped in the next minute and some objective chance that it will not. Now suppose that Heidi has bet that the coin will be flipped within the minute. But because of the posited indeterminacy, there are familiar arguments to the conclusion that both "The coin will be flipped within a minute" and "Heidi will win the bet" lack witnesses for their truth or falsity. The classical connectives are truth functional and, so, have no defined output for cases involving gappy sentences. Hence

It's not the case that the coin will be flipped within a minute ∨ Heidi will win the bet

is neither true or false. Nevertheless, If the coin will be fliped within a minute, Heidi will win the bet. So the indicative does not entail its material counterpart.

I go on to provide a formal semantic framework for a conditional connective that validates Modus Ponens along with all the controversial principles involved in three of the best-known direct arguments yet respects counterexamples of the above sort. It follows, against orthodoxy, that the failure of entailment from the indicative to the material conditional is consistent with the validity of Modus Ponens, and one may coherently grant the key principles of the direct arguments without being forced to accept the equivalence of the the indicative and material conditional.

References

Briggs, R. (2012). Interventionist counterfactuals. *Philosophical Studies 160*, 139–166.

Fisher, T. (2016a). Causal counterfactuals are not interventionist counterfactuals. *Synthese*, DOI 10.1007/s11229–016–1183–0.

Fisher, T. (2016b). Counterlegal dependence and causation's arrows: Causal models for backtrackers and counterlegals. *Synthese*, DOI 10.1007/s11229–016–1189–7.

Pearl, J. (2000). *Causality: models, reasoning, and inferences* (1st ed.). Cambridge University Press.