

## Old and new problems for Woodward’s interventionist theory of causation

**Abstract:** Woodward’s (2003) interventionist theory of causation is based on the causal nets approach (Pearl, 2009; Spirtes, Glymour, & Scheines, 2000) which has its roots in turn in the investigation of Bayesian nets (Neapolitan, 1990; Pearl, 1997). The theory intends to reformulate most of the causal nets approach’s technically quite challenging content in a philosophically more transparent way by means of the notion of a possible intervention (cf. Glymour, 2004). To a certain extent, however, it also exceeds the causal nets approach: it is not, like the causal nets approach, based on the basic notion of direct causal connection (with respect to a variable set  $\mathbf{V}$ ) alone, but rather tries to illuminate this notion in terms of possible interventions. In a nutshell Woodward’s account states that  $X$  is a direct cause of  $Y$  (with respect to a variable set  $\mathbf{V}$ ) if and only if there is a possible intervention on  $X$  that, if accomplished, would change  $Y$ ’s probability distribution when all values of other variables in  $\mathbf{V}$  are held fixed by interventions.

Though Woodward’s interventionist theory of causation has become notably famous and has many supporters, there are also some well-known problems it has to cope with (cf. Woodward, 2008). It is, for example, quite unclear what it means for an intervention to be a “possible” intervention. Since there are many causes one cannot intervene on—just take the gravitational constant or any closed physical system as an example—‘possible’ seems not to be intended to be read as ‘physically possible’. In this paper, however, we are mainly interested in the application of Woodward’s theory of causation, and thus, have to interpret possible interventions at least as “physically” possible. Another well-known objection to Woodward’s theory of causation is that it is circular because the notion of an intervention it uses is in turn explicated by means of causal notions. Woodward counters that the circle is not a vicious one: while the application of his account definitely requires some causal knowledge, it does not presuppose the causal knowledge its application should produce. In this paper we stay neutral on the question of whether Woodward’s reply to the circularity objection is satisfying or not. Instead we discuss some old and present some new problems with Woodward’s approach which arise independently of how this question may be answered.

After introducing some formal preliminaries, we present and explain Woodward’s interventionist theory of causation. We then show that there are some hypothetically possible systems Woodward’s theory cannot account for because its application would lead to false results in these scenarios. We present three sorts of systems of this kind: (i) systems containing variables one cannot intervene on at all or only by means of so-called “soft interventions”, (ii) causal chains with certain probabilistic properties, and (iii) systems containing deterministic causal chains of minimum length three. We highlight the conditions implemented within Woodward’s definition of an intervention variable and within his characterization of causal connection which lead to the exclusion of these sys-

tems and discuss possible ways out of the misery. By following our findings, step by step we develop an alternative version of Woodward's interventionist theory of causation which can omit the mentioned problems. We finally show that this alternative version we were forced to develop to deal with the observed problems does not have to be seen as an independent theory of causation, but can be proven as a theorem within the causal nets approach.

## References

- Glymour, C. (2004). [Critical notice to the book *Making things happen: A theory of causal explanation*, by J. Woodward]. *British Journal for the Philosophy of Science*, 55, 779–790.
- Neapolitan, R. E. (1990). *Probabilistic reasoning in expert systems: Theory and algorithms*. New York: Wiley.
- Pearl, J. (1997). *Probabilistic reasoning in intelligent systems*. San Francisco: Morgan Kaufmann.
- Pearl, J. (2009). *Causality*. Cambridge: Cambridge University Press.
- Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge: The MIT Press.
- Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press.
- Woodward, J. (2008). Causation and manipulability. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2008 ed.). <http://plato.stanford.edu/archives/win2008/entries/causation-mani/>.