

Causal Stacks: A Theoretical Framework for Recurrent and Hierarchical Counterfactual Reasoning

Dominic Le

University of Toronto, Toronto, Ontario, Canada

Abstract

Counterfactual (CF) reasoning – the process of considering alternative events and their outcomes – plays a vital role in understanding causation in fields like cognitive psychology and philosophy of science. In this paper, I develop a theoretical framework of Structural Causal Stacks (SCS) that provides a conceptual structure to describe the relationships between related causal and counterfactual analyses. Then, I explore its useability for observing human reasoning by running 500 pilot simulations of causal stack agents. My simulation modelled Gerstenberg et al. (2013)'s experiment design, which measured whether people's judgements about the consequence of a counterfactual state changes depended on the order they considered the events. According to my preliminary results, the stack model replicated the asymmetry in backwards versus forward counterfactual reasoning, aligning with the established consensus in a cognitive psychology literature while extending a persistent explanation for successive analyses.