Title
Extending counterfactual reasoning models to unconstrained social explanations

Permalink
https://escholarship.org/uc/item/2785p63c

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Droop, Stephanie
Bramley, Neil R.

Publication Date
2023

Peer reviewed

Extending counterfactual reasoning models to unconstrained social explanations

Stephanie Droop
University of Edinburgh, Edinburgh, United Kingdom

Neil Bramley
University of Edinburgh, Edinburgh, Scotland, United Kingdom

Abstract

Counterfactual models perturb a situation model to measure the explanatory power of different causes. We modify one to explore how people explain others' behaviour in two online experiments. First, 90 UK-based adults rated the likelihood of various scenarios combining short biographies with characters' trajectories through a gridworld. Then 49 others saw each scenario and outcome, and verbally gave their best explanations for why the character went the way they did. Participants generated a range of explanations for even the most incongruous behaviour, citing factors outside the given information. Using the innovative features of crowdsourced parameters and free text responses, we show people anchor to real world dependency c.75% of the time, and stochastically sample the rest. We present an expanded version of the counterfactual effect size model (Quillien & Lucas 2022) which handle surprising behaviours and performs better than the existing model it is based on.