Reasoning about Causality in Games (Abstract Reprint)

Lewis Hammond[1], **James Fox**[1], **Tom Everitt**[2], **Ryan Carey**[1], **Alessandro Abate**[1], **Michael Wooldridge**[1]

[1]University of Oxford, United Kingdom
[2]DeepMind, United Kingdom

Abstract Reprint. This is an abstract reprint of a journal article by Hammond, Fox, Everitt, Carey, Abate, and Wooldridge (2023).

Abstract

Causal reasoning and game-theoretic reasoning are fundamental topics in artificial intelligence, among many other disciplines: this paper is concerned with their intersection. Despite their importance, a formal framework that supports both these forms of reasoning has, until now, been lacking. We offer a solution in the form of (structural) causal games, which can be seen as extending Pearl's causal hierarchy to the game-theoretic domain, or as extending Koller and Milch's multi-agent influence diagrams to the causal domain. We then consider three key questions: i) How can the (causal) dependencies in games – either between variables, or between strategies – be modelled in a uniform, principled manner?

ii) How may causal queries be computed in causal games, and what assumptions does this require?

iii) How do causal games compare to existing formalisms?

To address question i), we introduce mechanised games, which encode dependencies between agents' decision rules and the distributions governing the game. In response to question ii), we present definitions of predictions, interventions, and counterfactuals, and discuss the assumptions required for each. Regarding question iii), we describe correspondences between causal games and other formalisms, and explain how causal games can be used to answer queries that other causal or game-theoretic models do not support. Finally, we highlight possible applications of causal games, aided by an extensive open-source Python library.

References

Hammond, L.; Fox, J.; Everitt, T.; Carey, R.; Abate, A.; and Wooldridge, M. 2023. Reasoning about causality in games. *Artificial Intelligence*, 320: 103919.