Causes and Strategies in Multiagent Systems
Extended Abstract

Sylvia S. Kerkhove[1][0009-0000-7153-3830], Natasha A. Alechina[2,1][0000-0003-3306-9891], and Mehdi M. Dastani[1][0000-0002-4641-4087]

[1] Utrecht University, Utrecht, The Netherlands
`{s.s.kerkhove,m.m.dastania}@uu.nl`
[2] Open University, Heerlen, The Netherlands
`natasha.alechina@ou.nl`

Abstract. Causality plays an important role in daily processes, human reasoning, and artificial intelligence. There has however not been much research on causality in multi-agent strategic settings. In this work, we introduce a systematic way to build a multi-agent system model, represented as a concurrent game structure, for a given structural causal model. In the obtained so-called causal concurrent game structure, transitions correspond to interventions on agent variables of the given causal model. The Halpern and Pearl framework of causality is used to determine the effects of a certain value for an agent variable on other variables. The causal concurrent game structure allows us to analyse and reason about causal effects of agents' strategic decisions. We formally investigate the relation between causal concurrent game structures and the original structural causal models.
This is an extended abstract of our paper Causes and Strategies in Multiagent Systems, published at AAMAS 2025 [13].

Keywords: Causality · Multi-Agent Systems · Strategic Behaviour

1 Introduction

Causality plays an important role in Artificial Intelligence [15, 12]. A specific type of causality, called 'actual causality', concerns causal relations between concrete events (e.g. throwing a specific rock shatters a specific bottle) [12]. There are multiple definitions of actual causality (see e.g. [12, 11, 9] and [6]), but most approaches (like [9] and [6]) use Pearl's [15] structural model framework. In this structural model framework, the world is modelled through variables, which are divided in exogenous and endogenous variables. The former are variables whose values are determined by causes outside of the model and the latter are variables whose values are determined by the variables inside the model (both exogenous and endogenous). The functional dependencies between variables are formalised through structural equations.

While causal models can in principle depict multi-agent systems by making a distinction between agent and environment events, they are less appropriate for reasoning about the abilities and strategies of agents. Concurrent game structures (CGS) have been proposed to reason about agent interactions and strategies [2]. These structures are graphs where nodes correspond to states of the world and edges, labelled with agents' actions, correspond to state transitions [4, 10]. In deterministic settings, an agent strategy specifies the actions to take by the agent.

Let us introduce an example of a causal model. Consider a semi-autonomous vehicle controlled jointly by a human driver and an automatic driving assistance system. This driving assistance system is in turn supported by an obstacle detection system that signals to the driving assistant whether there is an obstacle in front of the vehicle. Both the human driver and the driving assistant control the forward movement of the vehicle, though the human driver can always take full control. In a scenario where there is an obstacle in front of the car, the obstacle causes the obstacle detection system to send a signal to the driving assistant. If the human driver is in a distracted state, this signal causes the driving assistant to avoid an accident. This scenario can be described as a causal system, but can also be viewed as a multi-agent system where the obstacle detection system, the driving assistant and the human driver are all seen as agents that make decisions based on their state observations.

The fundamental relationship between structural causal models (SCMs) and multi-agent system models manifests itself in modelling phenomena such as responsibility for realising a certain outcome by a group of agents. In the literature of multi-agent systems, both structural causal models and CGS are used to define the responsibility of a group of agents for an outcome [7, 16]. Agents in a structural causal model are seen as responsible for an outcome if they have caused it [7]. On the other hand, in a CGS a coalition of agents is deemed responsible for an outcome if they had a strategy to prevent it [16]. By establishing the relationship between structural causal models and CGS, different modeling approaches to multi-agent phenomena (e.g., responsibility) can be compared and unified.

2 Causal Concurrent Game Structure

In our paper Causes and Strategies in Multiagent Systems [13], we aimed to establish a formal relationship between structural causal models and concurrent game structures by constructing a CGS for a given structural causal model such that a group of agents is an actual cause for an outcome in the causal model, if and only if this group had a strategy in the constructed CGS to prevent the outcome, provided the other agents act as prescribed by the causal model. The CGS is built by distinguishing between agent and environment variables. We consider the values of an agent variable as the possible actions of the agent and interventions as the agents' decisions, agents take their actions at a point corresponding to their position in the structural causal model. The causal CGS is defined in such a way that the leaf-states correspond to interventions on the original SCM. In [13], we work out an example for how this works for the semi-autonomous vehicle example we discussed in the introduction.

3 Discussion

In [13], we have only looked at deterministic and recursive causal models to define the causal CGS. However, causal relations are often probabilistic and cyclic in many practical use cases. Modelling such cases requires probabilistic and non-recursive causal models to, for example, capture the mutual dependencies between agents. In order to deal with probabilities, we will have to either employ probabilistic CGS, or use another type of model (e.g. Markov games). Moreover, allowing cyclic dependencies would make the evaluation of the states difficult, as the variable values would depend on each other. We think that this could possibly be dealt with by adding a temporal component to the model, but this needs more research.

Another direction of future work would be to use this framework to compare different approaches to defining responsibility in multi-agent settings. Some existing works define responsibility based on causal relations between agents and an outcome (like [1, 7, 8] and [5]), while other work is based on whether agents had a strategy to avoid the outcome (like [3] and [16]). The definition of causal CGS might help to combine both directions of research. Moreover, we can also look at how our approach compares to rule-based approaches to causality. Since Lorini's [14] work shows a correspondence between his rule-based framework for causal reasoning and the structural equations framework, it seems possible that his framework can also be shown to have a connection to our causal CGS.

We believe that our framework will be beneficial for supporting causal inference in multi-agent systems, for example, for reasoning and attributing responsibility for certain outcomes to groups of agents.

This research could be used in multi-agent systems with a clear causal structure. Examples of this are traffic control environments, like planes that cannot land when another is departing, trains that cannot travel over the same track at the same time, or traffic lights on a junction that cannot all turn to green at the same time. Other applications could be in the analysis of multi-player games, after all, players could cause other players to make a certain move, or energy management systems, where supply and demand of electricity influence each other. This research could help making decisions in these situations, or after something has gone wrong, it can help with the attribution of responsibility.

Acknowledgments. This study is part of the CAUSES project (KIVI.2019.004) of the research programme Responsible Use of Artificial Intelligence which is financed by the Dutch Research Council (NWO) and ProRail.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alechina, N., Halpern, J.Y., Logan, B.: Causality, responsibility and blame in team plans. In: Proceedings of the 16th Conference on Autonomous Agents and Multi-Agent Systems. pp. 1091–1099. ACM (May 2017)
2. Alur, R., Henzinger, T.A., Kupferman, O.: Alternating-time temporal logic. Journal of the ACM **49**(5), 672–713 (Sep 2002). https://doi.org/10.1145/585265.585270
3. Baier, C., Funke, F., Majumdar, R.: A game-theoretic account of responsibility allocation. Tech. Rep. arXiv:2105.09129 (2021)
4. Baier, C., Katoen, J.P.: Principles of model checking. MIT press (2008)
5. Beckers, S.: Moral responsibility for ai systems. In: Oh, A., Neumann, T., Globerson, A., Saenko, K., Hardt, M., Levine, S. (eds.) Advances in Neural Information Processing Systems. vol. 36, pp. 4295–4308. Curran Associates, Inc. (2023)
6. Beckers, S., Vennekens, J.: A principled approach to defining actual causation. Synthese **195**(2), 835–862 (Feb 2018). https://doi.org/10.1007/s11229-016-1247-1
7. Chockler, H., Halpern, J.Y.: Responsibility and blame: A structural-model approach. Journal of Artificial Intelligence Research **22**, 93–115 (2004)
8. Friedenberg, M., Halpern, J.Y.: Blameworthiness in multi-agent settings. Proceedings of the AAAI Conference on Artificial Intelligence **33**(01), 525–532 (Jul 2019). https://doi.org/10.1609/aaai.v33i01.3301525
9. Gladyshev, M., Alechina, N., Dastani, M., Doder, D., Logan, B.: Dynamic causality. In: Proceedings of the 26th European Conference on Artificial Intelligence. pp. 867–874 (2023). https://doi.org/10.3233/FAIA230355, https://ebooks.iospress.nl/volumearticle/64287
10. Gorrieri, R.: Process algebras for Petri nets: the alphabetization of distributed systems. Springer (2017)
11. Hall, N.: Structural equations and causation. Philosophical Studies **132**(1), 109–136 (Jan 2007). https://doi.org/10.1007/s11098-006-9057-9
12. Halpern, J.Y.: Actual causality. MIT Press (2016)
13. Kerkhove, S.S., Alechina, N., Dastani, M.: Causes and strategies in multiagent systems. In: 24th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2025. pp. 1098–1106. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS) (2025)
14. Lorini, E.: A rule-based modal view of causal reasoning. In: Elkind, E. (ed.) Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23. pp. 3286–3295. International Joint Conferences on Artificial Intelligence Organization (8 2023). https://doi.org/10.24963/ijcai.2023/366, main Track
15. Pearl, J.: Causal diagrams for empirical research. Biometrika **82**(4), 669–688 (1995)
16. Yazdanpanah, V., Dastani, M., Alechina, N., Logan, B., Jamroga, W.: Strategic responsibility under imperfect information. In: Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems AAMAS 2019. pp. 592–600. IFAAMAS (2019)