Causal Influence Detection for Human Robot Interaction

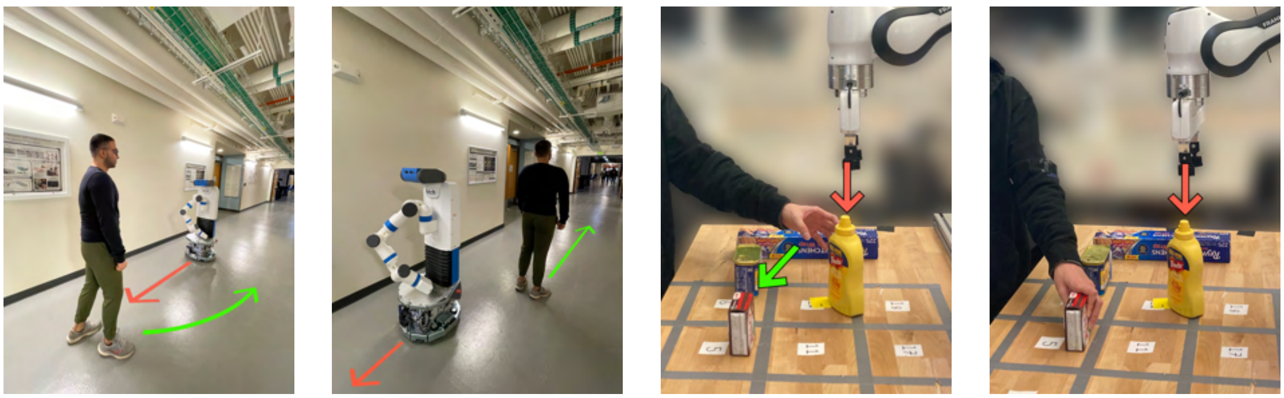
Yi-Shiuan Tung
yi-shiuan.tung@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

Himanshu Gupta
himanshu.gupta@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

Wei Jiang
wei.jiang@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

Bradley Hayes
bradley.hayes@colorado.edu
University of Colorado Boulder
Boulder, CO, USA

Alessandro Roncone
alessandro.roncone@colorado.edu
University of Colorado Boulder
Lab0 Inc.
Boulder, CO, USA



(a) (b) (c) (d)

Figure 1: The robot only influences the human's actions in certain scenarios. (a) The robot can choose to walk towards the human or move to its left to avoid the human. Both actions affect the human's path. (b) The robot's actions no longer affect the human's actions. (c) The human and the robot are clearing the table together. The robot reaches for the yellow mustard bottle, and the human may reach for another item or continue to pick up the mustard. The robot influences the human's actions. (d) The human picked up the jello box. The robot's actions in this state do not influence the human's next action: move the jello.

ABSTRACT

In human-robot and multi-agent interaction, the ego agent models the influence of its actions on the actions of the other agents to better anticipate what the other agents will do next, facilitating effective collaboration and enhanced safety. Prior work assumes that the ego agent has influence in all states when in reality the influence is only present in a subset of the scenarios. In this work, we propose to detect causal influence by measuring the mutual information of the ego agent's actions and the other agent's actions. We evaluate our approach in a simulated pedestrian navigation and a collaborative cooking game. Our results show that causal influence detection is a promising approach, yet it may yield low accuracy in situations where there is insufficient data.

HRI '24 Causal-HRI Workshop, March, 2024, Boulder, CO.

CCS CONCEPTS

• **Mathematics of computing** → *Information theory*; • **Computing methodologies** → **Causal reasoning and diagnostics**; *Multi-agent planning*.

KEYWORDS

multi-agent systems, human-robot interaction, causal influence detection

ACM Reference Format:
Yi-Shiuan Tung, Himanshu Gupta, Wei Jiang, Bradley Hayes, and Alessandro Roncone. 2024. Causal Influence Detection for Human Robot Interaction. In *HRI '24: Workshop on Causal Learning for Human-Robot Interaction (Causal -HRI), HRI2024.* ACM, New York, NY, USA, 5 pages.

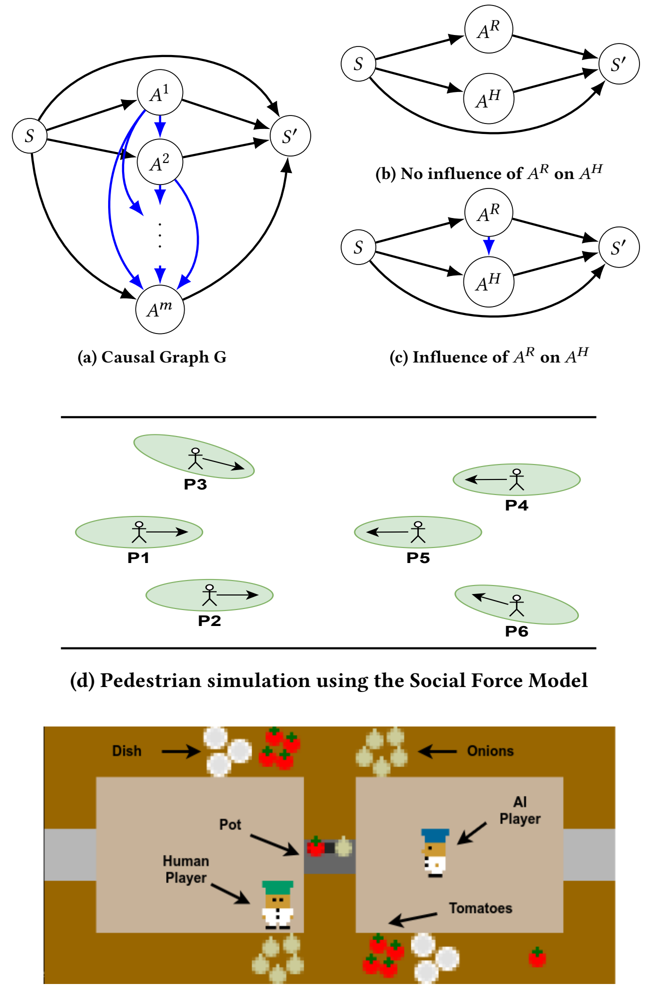
1 INTRODUCTION

In human-robot interaction, the standard method involves the robot inferring the human's intention and subsequently planning an action to collaborate with the human [8, 17]. In practice, the human is also inferring the robot's intentions and modifying their actions in response to the robot's actions. For example, imagine that you are walking down a hallway where a robot approaches from the opposite direction (Fig. 1a). If the robot immediately moves to the side, you will continue to walk straight since you will not collide with the robot. However, if the robot walks directly towards you, you may suspect that the robot has faulty sensors or is in a hurry and move out of the robot's way.

To facilitate fluent interaction, the robot has to learn how its actions will influence the human's actions, otherwise, the robot's predictions of human intentions may be inaccurate, resulting in suboptimal decision-making. Prior work on robots accounting for the impact of their actions on the other agents' actions assumes that this influence occurs at fixed intervals [14, 16] or the entirety of the interaction [6, 7]. In reality, the robot is only able to influence other agent's actions in certain scenarios. In the hallway navigation example, once the robot passes the human, the robot doesn't alter the human's path by changing lanes (Fig. 1b).

Learning when and how a robot can influence the actions of other agents is beneficial for both learning and exploration. A major challenge of multi-agent reinforcement learning (MARL) is the non-stationarity of the environment due to the changing policies of other agents. By taking into account the learning of other agents and the cause-and-effect relationships of agent interactions, effective agents can quickly adapt to non-stationary behaviors [6]. In the context of reinforcement learning (RL), empowerment refers to an agent's ability to influence the future states given the current state and the actions it can take [1]. In MARL, our approach can empower agents to explore states with high influence over other agents' actions, leading to more effective cooperation [13].

In this work, we propose to use causal influence detection to determine when a robot's actions influence the human's actions by finding the existence of directed edges in a causal graph. We extend the work by Seitzer et al. [12] which uses a measure of causal influence based on conditional mutual information to detect states where a single agent can influence the outcome of the next state. We compute the mutual information between the human and the robot's actions and infer the presence of influence when this value is above a predetermined threshold. To evaluate our approach, we simulate pedestrian navigation using the Social Force Model (SFM) [5] and compare the causal influence predictions with the ground truth derived from the simulation. Furthermore, we train a deep neural network on human-human and human-AI gameplay from a collaborative cooking game to predict causal influence and evaluate the model on held out test data. Our results show that causal influence detection is a promising approach for identifying when the actions of a robotic agent influence the actions of other agents and generalizes well to unseen test data. However, it can be inaccurate when the data does not sufficiently cover the full support of the action distribution.



(a) Causal Graph G

(b) No influence of A^R on A^H

(c) Influence of A^R on A^H

(d) Pedestrian simulation using the Social Force Model

(e) Collaborative cooking game

Figure 2: Causal graphical model of a single state transition. (a) Each agent's actions potentially influence the other agents' actions. Our goal is to detect this influence by finding the existence of the blue edges. (b) In a dyadic human-robot interaction scenario, if the robot's action A^R does not influence the human's action A^H in state S, there is no arrow from A^R to A^H. (c) The blue arrow indicates that A^R has influence on A^H. (d) The agent's actions have no influence on the actions of the other agents because the social repulsive force represented by the green ellipses do not overlap. (e) If the AI player moves up to pick up the onion, the human player may move up to pick up the dish instead. A^R has influence on A^H in this state.

2 RELATED WORKS

Causal inference provides a framework for determining cause and effect relationships from data. The causal relationships can be described by a Causal Graphical Model (CGM), which is a directed acyclic graph where an edge from node X to Y indicates that the value of Y depends on the value of X. Given a CGM, one can explicitly reason about interventions by using the do operator to force a variable X to take on value x [11]. In MARL, Foerster et al. [3] uses counterfactuals to address the challenges of multi-agent credit assignment by marginalizing out a single agent's actions. Meganck et al. [10] introduces multi-agent causal models and proposes a decentralized method for learning the global model. Other works embed a causal model in the multi-agent framework and apply it to applications such as healthcare, advertising, and traffic control [4].

In contrast, our work uses causal discovery to learn a CGM that determines whether a robot's actions affects the human's actions in a specific state.

In human-robot interaction and MARL, prior work models the influence of the robot's actions on the other agents' actions to improve agent learning and adaptation. Kim et al. [7] proposes an algorithm based on policy gradient and bellman update that allows an agent to influence the long term stationary distribution of the system. Kim et al. [6] introduces a policy gradient that includes an extra optimization term for how an agent's current actions affects the other agents' policy updates. Xie et al. [16] uses an autoencoder to predict the human's latent strategy and its dynamics given the state and the robot's actions. Wang et al. [14] trains the robotic agent with an additional reward of stabilizing the human's latent strategy, predicted by an autoencoder, to address the challenge of non-stationarity in MARL. Several works explored recursive reasoning by modeling how the other agents would react to the ego agent's future behaviors [9, 15]. These works assume that the agent influences the actions of the other agents across all states whereas our paper challenges this assumption. Instead, we employ causal influence detection to identify specific states where the robot has the potential to influence the actions of other agents.

3 METHODS

3.1 Causal Graphical Models

To model multi-agent Markov decision processes, we use the Causal Graphical Model (CGM) which consists of a set of random variables $V = \{S, A^1, A^2, ..., A^m, S'\}$, a graph G, and conditional distributions $p(v_i|Pa(V_i))$ where $Pa(V_i)$ is the set of parents of V_i (Fig. 2a). The current state is represented by S, A^i is the action of the ith agent, and S' is the next state. The joint distribution of V, denoted as P_V, can be factorized as $p(v_1, ..., v_{|V|}) = \prod_{i=1}^{|V|} p(v_i|Pa(V_i))$. In this paper, we are interested in detecting whether an edge exists between the agent's actions. For dyadic human-robot teaming, A^R has an edge to A^H if the robot's action influences the human's actions in state S (Fig. 2c).

3.2 Causal Influence Detection

We extend Seitzer et al. [12], which proposes causal influence detection through computing the mutual information, to detect action influence in a multi-agent setting. The mutual information $I(A^R; A^H)$ is a measure of dependence between random variables A^R and A^H and is zero for no dependence. $D_{KL}(\cdot||\cdot)$ denotes the KL divergence. The mutual information I is defined as

$$\begin{aligned} I(A^R; A^H|S = s) &= D_{KL}(p(A^R, A^H|s)||p(A^R|s) \otimes p(A^H|s)) \\ &= \sum_{a^R \in A^R} \sum_{a^H \in A^H} p(a^R, a^H|s) log(\frac{p(a^R, a^H|s)}{p(a^R|s)p(a^H|s)}) \end{aligned} \tag{1}$$

We estimate the probability distributions from a dataset via visitation counts. $p(a^R, a^H|s)$ is the ratio of $N(a^R, a^H, s)$ to $N(s)$ where the numerator is the number of times actions a^R and a^H are executed in state s and the denominator is the total number of times state s is visited across all the data points. $p(a^R|s)$ and $p(a^H|s)$ are the marginal probability distributions and are computed by summing $p(a^R, a^H|s)$ over all possible a^H and a^R respectively.

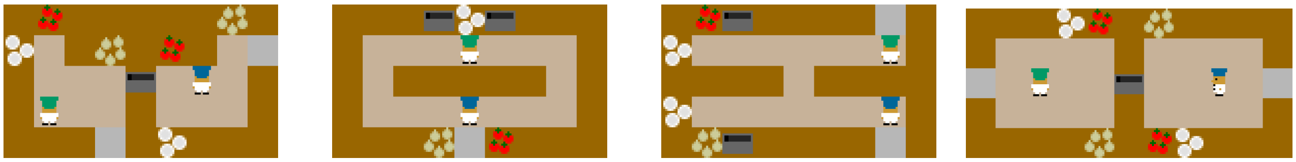
4 EXPERIMENTS

In our experiments, we aim to answer two questions: 1) Can mutual information reliably detect states where the ego agent influences the other agent's actions? 2) How well does a neural network learn causal influence detection and can it generalize to unseen states? To answer the first question, we simulate pedestrians walking down a hallway using the Social Force Model (SFM) [5]. From observing the positions and actions of the pedestrians, we compute the mutual information between the ego agent's actions and the neighboring pedestrians' actions in each state and infer causal influence when the value is above a predetermined threshold. To answer the second question, we use human-human and human-AI gameplay data of a collaborative cooking game called Overcooked [2]. We train a neural network to learn causal influence determined by computing mutual information of the robot and the human's actions, and we evaluate the model's performance on held out test data. The neural network consists of a single hidden layer comprising 64 units with the ReLU (Rectified Linear Unit) activation function.

4.1 Social Navigation

We use the SFM introduced in [5] to simulate pedestrians walking down a hallway. According to SFM, a pedestrian's motion is influenced by its tendency to stay near the shortest path to its goal, the motion of other pedestrians in its vicinity, the attraction it has towards familiar objects, and the repulsion it has towards walls and other obstacles in the environment. The territorial effect of each pedestrian induces a repulsive force in the shape of an ellipse along the direction of its motion. This force is a monotonic decreasing function of the distance from that pedestrian.

We ran 10 trials with varying number of pedestrians randomly initialized on each side of the hallway traversing towards the opposite side. A single entry in our dataset is of the form (S, A^1, A^2, L) which denotes the joint state, the action of the ego agent, the action of a neighboring agent, and the label indicating influence, respectively. Let's assume pedestrian H^1 is the ego agent. We find the closest human to the ego agent, H^2. We then compute the distance vector from H^1 to H^2, denoted as $\vec{a}$. We also find the closest pedestrian to H^2, denoted as H^j, and compute the distance vector from H^2 to H^j, denoted as $\vec{b}$. We include the orientation θ of H^1 and H^2 so the model can differentiate if the agents are walking towards or away from each other. The joint state is then defined as $S = (\vec{a}, \vec{b}, \theta_{H^1}, \theta_{H^2})$. The action for each pedestrian is obtained using SFM and is a two-dimensional vector that consists of its change in orientation, $\Delta\theta$, and change in speed, Δv. The action for the ego agent is defined as $A^1 = (\Delta\theta^1, \Delta v^1)$ and that of H^2 is defined as $A^2 = (\Delta\theta^2, \Delta v^2)$. For tractability, we discretize the joint state and the action for both H^1 and H^2. The label L is marked as 1 if H^1 falls in the repulsive force ellipse of H^2, otherwise 0. The tuple (S, A^1, A^2, L) is similarly computed with the second closest human to the ego agent. This is repeated for every time step of the trial and by treating every human as the ego agent.



(a) Asymmetric Advantages (b) Counter Circult (c) Cramped Corridor (d) Soup Coordination

Figure 3: The Overcooked layouts that we used to analyze causal influence detection. The diverse layouts offer different coordination strategies and thus varying levels of action influence.

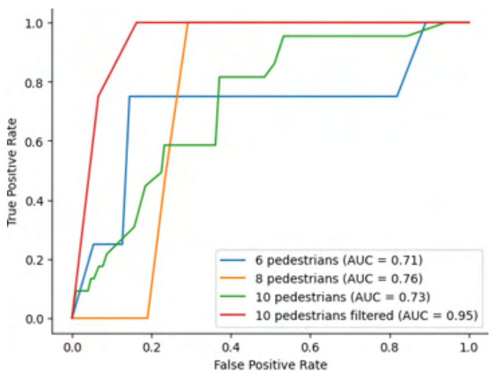


Figure 4: We simulate pedestrian navigation using SFM and use Eqn. 1 to detect causal influence. We compare the predictions to the ground truth acquired from the SFM (influence is present if the social repulsive forces overlap) and plot the receiver operating characteristic curve. The model performs similarly for different number of pedestrians but improves when we filter out states with few interaction data.

4.2 Collaborative Cooking Task

Overcooked is a collaborative game where players place ingredients into a pot and deliver the cooked dish to a serving station. It is a challenging collaboration game due to the variety of coordination strategies available, requiring players to adapt to one another's actions to achieve high scores. We use the grid world implementation and the human-human and human-AI gameplay data collected by Carroll et al. [2]. The dataset contains different environment layouts, facilitating the emergence of diverse coordination strategies. We use the train test splits provided by Carroll et al. [2] [1].

To get the ground truth binary labels, we compute mutual information of the human and the robot's actions and infer causal influence if the value is above 0.5.

5 PRELIMINARY RESULTS

5.1 Social Navigation

Figure 4 plots the receiver operating characteristic curve (ROC) when simulating different number of pedestrians. The ROC curve shows the performance of our causal influence detection at various classification thresholds. The model performs similarly for different number of pedestrians but improves when we filter out states with less than 8 interaction data points (labeled 10 pedestrians filtered). This demonstrates that causal influence detection is less accurate with few interaction data points, highlighting the importance of employing strategies that collect informative data for improved causal influence detection.

[1]https://github.com/HumanCompatibleAI/overcooked_ai

Table 1: Performance of model predictions of causal influence in different Overcooked layouts on test data

Metric	Asymmetric Advantages	Counter Circuit	Cramped Corridor	Soup Coordination
Precision	0.955	0.913	0.929	0.919
Recall	0.947	0.962	0.934	0.877
F1 Score	0.951	0.936	0.931	0.898

5.2 Collaborative Cooking Task

Table 1 shows the precision, recall, and F1 scores for the trained neural network when evaluated on the test data for different Overcooked environments. The corresponding layouts are shown in Figure 3. The trained model is able to generalize causal influence detection to states unseen during training.

6 CONCLUSION

In this work, we propose using causal influence detection via a measure of mutual information to detect an ego agent's influence on the actions of the other agents. We simulated pedestrian navigation and showed that computing mutual information recovers some of the states where influence is present but suffers from lack of action diversity in the data. Through the analysis of human and AI gameplay in Overcooked, we demonstrate the ability of a neural network to predict causal influence on unseen test data.

Future work should investigate how causal influence detection can be improved by incorporating it as an objective in the robot's decision making process. Furthermore, we want to investigate how the knowledge of this interaction influence can enable a robotic agent to better adapt and coordinate with humans and other agents.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research under Grant N00014-22-1-2482 and the Army Research Laboratory under Grant W911NF-21-2-0126.

REFERENCES

[1] Arthur Aubret, Laetitia Matignon, and Salima Hassas. 2019. A survey on intrinsic motivation in reinforcement learning. *arXiv preprint arXiv:1908.06976* (2019).

[2] Micah Carroll, Rohin Shah, Mark K Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, and Anca Dragan. 2019. On the utility of learning about humans for human-ai coordination. *Advances in neural information processing systems* 32 (2019).

[3] Jakob Foerster, Gregory Farquhar, Triantafyllos Afouras, Nantas Nardelli, and Shimon Whiteson. 2018. Counterfactual multi-agent policy gradients. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32.

[4] St John Grimbly, Jonathan Shock, and Arnu Pretorius. 2021. Causal multi-agent reinforcement learning: Review and open problems. *arXiv preprint arXiv:2111.06721* (2021).

[5] Dirk Helbing and Peter Molnar. 1995. Social force model for pedestrian dynamics. *Physical review E* 51, 5 (1995), 4282.

[6] Dong Ki Kim, Miao Liu, Matthew D Riemer, Chuangchuang Sun, Marwa Abdulhai, Golnaz Habibi, Sebastian Lopez-Cot, Gerald Tesauro, and Jonathan How. 2021. A policy gradient algorithm for learning to learn in multiagent reinforcement learning. In *International Conference on Machine Learning*. PMLR, 5541–5550.

[7] Dong-Ki Kim, Matthew Riemer, Miao Liu, Jakob Foerster, Michael Everett, Chuangchuang Sun, Gerald Tesauro, and Jonathan P How. 2022. Influencing long-term behavior in multiagent reinforcement learning. *Advances in Neural Information Processing Systems* 35 (2022), 18808–18821.

[8] Dylan P Losey, Craig G McDonald, Edoardo Battaglia, and Marcia K O'Malley. 2018. A review of intent detection, arbitration, and communication aspects of shared control for physical human–robot interaction. *Applied Mechanics Reviews* 70, 1 (2018), 010804.

[9] Xiaobai Ma, David Isele, Jayesh K Gupta, Kikuo Fujimura, and Mykel J Kochenderfer. 2022. Recursive Reasoning Graph for Multi-Agent Reinforcement Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 7664–7671.

[10] S. Meganck, S. Maes, B. Manderick, and P. Leray. 2005. Distributed learning of multi-agent causal models. In *IEEE/WIC/ACM International Conference on Intelligent Agent Technology*. 285–288. https://doi.org/10.1109/IAT.2005.66

[11] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. 2017. *Elements of causal inference: foundations and learning algorithms*. The MIT Press.

[12] Maximilian Seitzer, Bernhard Schölkopf, and Georg Martius. 2021. Causal influence detection for improving efficiency in reinforcement learning. *Advances in Neural Information Processing Systems* 34 (2021), 22905–22918.

[13] Tessa van der Heiden, Christoph Salge, Efstratios Gavves, and Herke van Hoof. 2020. Robust multi-agent reinforcement learning with social empowerment for coordination and communication. *arXiv preprint arXiv:2012.08255* (2020).

[14] Woodrow Zhouyuan Wang, Andy Shih, Annie Xie, and Dorsa Sadigh. 2022. Influencing towards stable multi-agent interactions. In *Conference on robot learning*. PMLR, 1132–1143.

[15] Ying Wen, Yaodong Yang, Rui Luo, Jun Wang, and Wei Pan. 2019. Probabilistic recursive reasoning for multi-agent reinforcement learning. *arXiv preprint arXiv:1901.09207* (2019).

[16] Annie Xie, Dylan Losey, Ryan Tolsma, Chelsea Finn, and Dorsa Sadigh. 2021. Learning latent representations to influence multi-agent interaction. In *Conference on robot learning*. PMLR, 575–588.

[17] Canjun Yang, Yuanchao Zhu, and Yanhu Chen. 2021. A review of human–machine cooperation in the robotics domain. *IEEE Transactions on Human-Machine Systems* 52, 1 (2021), 12–25.