Causal Reflection with Language Models

Abi Aryan
Abide AI

Zac Yung-Chun Liu
Abide AI

Abstract

Large Language Models (LLMs) and traditional Reinforcement Learning (RL) agents lack robust causal reasoning, often relying on spurious correlations. We introduce Causal Reflection, a framework that moves beyond simple reward optimization to build dynamic causal models of an environment. Our approach features a temporal, action-based causal function that models state, action, time, and perturbation to capture delayed and nonlinear effects. We also define a formal `Reflect` mechanism that identifies mismatches between predicted and observed outcomes, generating causal hypotheses to revise the agent's internal model. Within this architecture, LLMs are not black-box reasoners but structured interpreters, translating formal causal outputs into natural language explanations. This work lays the theoretical groundwork for agents that can adapt, self-correct, and communicate causal understanding.

1 Introduction

The exponential growth in artificial intelligence capabilities has intensified the need for systems that understand not just what happens, but why. Traditional reinforcement learning (RL) paradigms, while successful in maximizing reward signals, fundamentally lack the capacity to model the temporal cause-effect relationships that govern dynamic systems [Kiciman et al., 2023, Seitzer et al., 2021]. This limitation becomes especially pronounced when agents must adapt to changing environments, explain their decisions, or transfer learned behaviors across domains, particularly in business and enterprise settings where resilient decision systems are critical. Similarly, while large language models (LLMs) excel at knowledge synthesis and reasoning over static information, they also lack an inherent understanding of causality in temporal contexts [Jiao et al., 2024, Du et al., 2017]. Despite their promise, the integration of LLMs with causal reasoning for decision-making over time remains largely unexplored.

Addressing these challenges requires a paradigm shift away from correlation-based models. To this end, recent advances in causal inference offer a promising path forward by replacing them with interpretable causal frameworks designed to capture stable and invariant relationships [Deng et al., 2023]. Within reinforcement learning, this approach directly confronts critical limitations, preventing agents from succumbing to spurious correlations and enabling them to generalize beyond their training environments [Wan et al., 2024]. Several studies have begun exploring causal reinforcement learning by incorporating structural causal graphs to improve sample efficiency and robustness [Liu et al., 2025, Peters et al., 2017]. However, these methods typically assume fixed causal structures and thus fail to capture the dynamic nature of cause-effect relationships that change over time [He et al., 2025]. Moreover, the current integration of LLM reasoning capabilities with RL environments remains superficial, lacking a principled framework for state understanding. This gap is particularly evident in scenarios requiring agents to reason about delayed effects, temporal dependencies, and the evolution of causal mechanisms, i.e. the capabilities essential for real-world deployment where causality unfolds across extended time horizons.

This paper introduces a novel framework that addresses these limitations through three key contributions. First, we present a temporal, action-based causal function, enabling the capture of both linear

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Efficient Reasoning.

and nonlinear causal relationships in dynamic systems. Second, we demonstrate how this framework extends beyond traditional reinforcement learning paradigms. Third, we outline how LLMs can integrate this causal schema for enhanced generative and predictive tasks. In sum, our approach represents a fundamental shift toward causal reflection in agents for more interpretable, robust, and generalizable artificial intelligence systems.

2 Background and Related Work

The convergence of causal inference and reinforcement learning is an emerging paradigm that addresses fundamental limitations in traditional sequential decision-making systems. This section provides a systematic review of the foundational concepts and methodological approaches that inform our framework.

2.1 Causal Inference in Reinforcement Learning

Integrating causal reasoning into reinforcement learning (RL) has emerged as a critical research direction, prompted by the realization that traditional RL agents often rely on spurious correlations rather than true cause-effect relationships [Deng et al., 2023]. Causal Reinforcement Learning (CRL) addresses this limitation by embedding causal understanding directly into the learning process [Deng et al., 2023].

CRL leverages formalisms like Structural Causal Models (SCMs) to represent the underlying mechanics of an environment, moving beyond surface-level associations. Grounded in Pearl's Causal Hierarchy, this approach enables agents to reason not only associatively, but also interventionally and counterfactually. As a result, agents gain a rudimentary form of self-reflection, evaluating alternative histories to improve generalization, sample efficiency, and explainability [Bareinboim et al., 2021].

In practice, current CRL methodologies apply causal principles to improve policy learning. For instance, some approaches integrate causal discovery with reinforcement learning by using causal Dynamic Bayesian Networks to model state–action–reward relationships, enabling more informed action selection [Méndez-Molina et al., 2022]. While promising, these techniques face challenges in scalability and are typically constrained to low-dimensional or simplified environments, exposing a significant gap in applying CRL to complex, high-dimensional settings [Méndez-Molina et al., 2022].

2.2 Temporal Causality Models

Modeling temporal causal relationships poses challenges distinct from static inference, requiring frameworks that capture how cause-effect dynamics evolve over time [Gkorgkolis et al., 2025, Calderon and Berman, 2024]. Temporal Structural Causal Models (TSCMs) extend traditional SCMs by explicitly representing time-varying causal structures and dependencies [Gkorgkolis et al., 2025]. Recent advances have also addressed non-stationarity in dynamic systems, for instance, Temporal Autoencoders for Causal Inference (TACI) enable the measurement of shifting causal interactions without the need for continual retraining [Calderon and Berman, 2024].

Despite these innovations, existing models often lack an internal mechanism for self-reflection. Self Relfection here refers to a model's ability to question and revise their own causal assumptions when confronted with structural breaks. This gap is the focus of our work.

These modeling challenges are further complicated by foundational questions about the nature of causality in temporal systems. For example, coupled chaotic systems have been shown to violate assumptions such as Granger causality, which presumes that causes precede effects in a linear, time-delayed fashion [Paluš et al., 2018]. Such findings highlight the importance of distinguishing between simple signal transfer and complex nonlinear interactions when designing robust causal models, especially those intended to operate across varying time scales [Paluš et al., 2018]. We believe, incorporating this understanding is essential for developing models that not only adapt to dynamic environments but also revise their causal assumptions in the face of unexpected structural changes.

2.3 Large Language Models in Causal Reasoning

The emergence of large language models (LLMs) has opened new frontiers for causal reasoning, with studies demonstrating state-of-the-art performance on benchmarks for pairwise causal discovery, counterfactual reasoning, and event causality [Kiciman et al., 2023]. Unlike traditional methods that rely on statistical relationships, LLMs leverage vast causal knowledge embedded in their training data to perform complex tasks, such as generating causal graphs from text and translating between formal and natural language representations [Kiciman et al., 2023, Gkountouras et al., 2024]. However, this capability is often shallow, typically aligned with Level-1 reasoning and lacks the self-reflective mechanisms needed to scrutinize its own decision-making [Wang and Shen, 2024, Ashwani et al., 2024, Chi et al., 2024]. Instead of understanding causal structures, it relies heavily on memorized patterns, which can lead to significant performance degradation in novel scenarios and unpredictable failures.

To address these limitations, recent work has explored self-correction, where models iteratively refine their outputs based on internal or external feedback. A parallel paradigm endows LLM agents with self-reflection capabilities, allowing them to verbally reflect on past failures to refine future plans and improve performance [Shinn et al., 2023]. Building on these ideas, a promising research direction is the integration of LLMs with structured causal frameworks [Liu et al., 2025, Gkountouras et al., 2024]. Hybrid approaches that combine language models with causal world models provide a systematic interface between natural language and formal causal structures, enabling more robust, causally-aware reasoning [Gkountouras et al., 2024]. These integrated systems have demonstrated superior performance compared to pure LLM-based methods, particularly for complex inference and long-horizon planning tasks [Gkountouras et al., 2024]. This informs our framework's approach, which aims to leverage LLM capabilities while maintaining rigorous causal foundations through structural modeling and counterfactual reasoning.

3 Theoretical Framework

This section formalizes the core technical contributions of our Causal Reflection framework. We begin by defining the fundamental concepts before presenting the Causal Reflection function and discussing its theoretical properties. The objective is to establish a rigorous mathematical foundation for modeling causality in dynamic and potentially non-stationary environments.

3.1 Core Concepts and Definitions

To ground our framework, let us first establish a formal vocabulary that extends traditional state-action representations to explicitly account for time and nonlinear influences.

- **State (S_t):** The state at time t, denoted as S_t, is a comprehensive vector representing the complete configuration of the environment. This representation is assumed to be sufficiently rich to capture all relevant variables describing the system at a given moment, which can implicitly include spatial, contextual, or environmental information.
- **Action (A_t):** An action, A_t, is an intervention performed by an agent or an external force on the system at time t. It represents a deliberate change intended to influence the system's trajectory.
- **Time (T_t):** Time, T_t, is a continuous or discrete variable that imposes a strict temporal ordering on states and actions. Its inclusion is critical for moving beyond mere association to inferring causal relationships, as causes are generally understood to precede their effects in physical systems.
- **Perturbation Factor (δ):** A key innovation of our framework is the introduction of a perturbation factor, δ. This term represents small, often unobserved or unexpected influences that can trigger nonlinear, disproportionate, or chaotic effects within the system. It formalizes the sensitivity to initial conditions, often described as the "butterfly effect," where a minor event can initiate a cascade of significant consequences.

The explicit modeling of δ provides a mechanism to address the critical challenge of non-stationarity and structural breaks in temporal causal models, a known limitation of existing approaches that often

assume fixed causal structures. By incorporating δ, our framework can represent environments where the underlying causal mechanisms themselves can evolve spuriously over time.

3.2 The Temporal Action-Based Causal Function

We formulated a causal function, C, that maps the state, action, time, and perturbation from current to a future state (Figure 1). We build this function progressively to illustrate its components.

Figure 1: General workflow of the Causal Reflection framework. An agent's causal model C makes a prediction, which is compared against the observed outcome. Discrepancies trigger the `Reflect` mechanism to generate hypotheses and update the model. The LLM translates the formal outputs into natural language.

First, we define a basic causal mapping that models direct, immediate effects:

$$C(S_t, A_t, T_t) \to S_{t+1} \tag{1}$$

This function expresses a simple causal relationship where an action A_t in state S_t at time T_t directly results in the subsequent state S_{t+1}. To account for real-world complexities where effects are often not instantaneous, we extend the function to incorporate temporal delays. By introducing a delay factor k, the function can model long-term causal relationships:

$$C(S_t, A_t, T_t, T_{t+k}) \to S_{t+k} \tag{2}$$

Here, the effect of an action taken at time T_t manifests at a future time T_{t+k}. This capability is the key for reasoning about systems with significant lags between cause and effect, a common feature in domains like economics, medicine, and climate science. And, finally, to capture the full dynamics of complex, non-stationary systems, we introduce the complete temporal action-based causal function, which integrates the perturbation factor δ:

$$C(S_t, A_t, T_t, \delta) \to S_{t+k} \tag{3}$$

This is the key for modeling systems where causality is not only delayed but also nonlinear and time-varying. For instance, in an ecological model, a small environmental perturbation (δ), such as a lightning strike, might nonlinearly amplify the effect of a routine event (Action A_t) on animal behavior (State S_{t+k}) in a manner that linear models cannot capture. We treat δ not as random noise but as a latent indicator of structural deviation capturing unmodeled influences such as adversarial actions, domain drift, or unobserved confounders.

3.3 Properties

The proposed causal function exhibits several key properties that distinguish it from traditional models.

1. **Nonlinear Causal Propagation:** The inclusion of δ allows the function to model how small causes can propagate through a system to produce disproportionate effects. This can be operationalized through specific functional forms, such as:

   $$S_{t+k} = S_t + f(A_t, T_t) \cdot e^{-\delta} \tag{4}$$

   In this example, $f(A_t, T_t)$ represents the standard dynamic effect of an action, while the term $e^{-\delta}$ acts as a nonlinear scaling factor based on the perturbation. This formulation explicitly models how unforeseen events can amplify or dampen causal effects.

2. **Temporal Dependency and Delayed Effects:** The explicit inclusion of T_t and the delay parameter k makes the framework inherently suited for reasoning about causality across extended time horizons. This directly addresses the challenges of modeling temporal dependencies noted in existing literature.

3. **Sensitivity to Initial Conditions:** By formalizing the role of δ, our framework is designed to be sensitive to the initial conditions of the system, a defining characteristic of complex and chaotic systems. This property is a critical differentiator from purely probabilistic or linear causal models, that may fail to predict state transitions in environments subject to structural breaks or emergent phenomena.

4 Framework Architecture

This section outlines the operational architecture of the Causal Reflection framework, detailing the system design and the novel integration schema for Large Language Models (LLMs).

The framework is designed as a modular system that separates causal inference from linguistic interpretation. An agent interacts with its environment, observing and collecting data to form a structured causal tuple (S_t, A_t, T_t, δ). This tuple serves as the input to the temporal action-based causal function, C. The function processes this input to generate a prediction about a future state, S_{t+k}, or a distribution over possible future states. This formal, symbolic output is then passed to the LLM component. The LLM's role is not to choose an action or optimize a policy, but to translate the abstract causal inference into an intelligible explanation or a set of testable hypotheses.

Our framework redefines the role of LLMs in causal reasoning systems. Instead of relying on the implicit, and often brittle, causal knowledge stored in an LLM's parameters, we utilize the LLM as a sophisticated generative inference engine.

- **Input:** The LLM receives the structured causal tuple (S_t, A_t, T_t, δ) and the corresponding predicted outcome(s) from the causal function C.
- **Processing:** The LLM leverages its vast linguistic capabilities to process this formal input. It is tasked with translating the symbolic representation of the cause-effect relationship into natural language.
- **Output:** The LLM generates outputs such as causal explanations (e.g., "The system is predicted to transition to state S_{t+k} because action A_t was performed, and its effect was significantly amplified by perturbation δ") or counterfactual hypotheses (e.g., "Had perturbation δ not occurred, the model predicts the system would have transitioned to state S'_{t+k} instead").

This integration schema directly addresses the widely recognized limitation of "shallow" or "Level-1" causal reasoning in LLMs. Current models often succeed by matching patterns seen during training but fail in novel scenarios because they lack a genuine model of causal mechanisms.

Our framework mitigates this by enforcing a separation of concerns: the rigorous causal logic resides entirely within our formal function C, while the LLM is responsible for the subsequent linguistic interpretation. The LLM is not asked an open-ended question like "What caused Y?"; it is given a constrained task: "Given that our formal model C predicts X causes Y under condition δ, generate an explanation." This grounds the LLM's output in an external, verifiable causal structure, providing a pathway toward more robust and trustworthy causal explanations from AI systems.

Additionally, we operationalize the concept of "self-reflection" through a formal mechanism that enables the agent to learn from experience by refining its internal causal model. This is centered around a `Reflect` function, which formalizes the process of analyzing discrepancies between predicted and observed outcomes. The mechanism is defined as:

$$S'_{t+k} = \text{Reflect}(S_t, A_t, T_t, \delta) \tag{5}$$

The Reflect function serves as the core of Causal Reflection. When an agent takes an action and the observed outcome differs from the one predicted outcome by its causal function C, the Reflect function is invoked. It uses the components of the causal tuple to generate hypotheses about the source of the discrepancy.

For example, the agent might reason: "My action A_t did not lead to the expected state. The Reflect function, analyzing the context, suggests this could be due to an unmodeled or misestimated perturbation δ. Hypothesis: a new type of perturbation is active under these conditions."

This approach represents a significant advance over existing self-reflection paradigms, such as the Reflexion framework [Shinn et al., 2023]. While Reflexion improves agent behavior through verbal

reinforcement and meta-cognitive feedback, its reflection process remains heuristic and unstructured. In contrast, Causal Reflection introduces a formal mechanism rooted in a causal model of the environment. Rather than interpreting failure as a vague signal to try a new plan, our agent detects structural mismatches between expected and observed outcomes, prompting it to generate specific, falsifiable causal hypotheses. This elevates reflection from intuition to inference, enabling the agent to function by explaining, testing, and revising its internal model of the world's causal structure. We formalize this process through a dedicated function, `Reflect`, which takes as input:

- The current state S_t, action A_t, time T_t, and estimated perturbation δ,
- The predicted outcome $\hat{S}_{t+k}$ generated by the causal model C,
- The observed outcome S_{t+k}^{obs}.

Letting

$$\epsilon = \text{Loss}(\hat{S}_{t+k}, S_{t+k}^{\text{obs}}) \tag{6}$$

the agent triggers reflection when ϵ exceeds a threshold τ, indicating a causal mismatch. The `Reflect` mechanism then initiates hypothesis generation:

$$H_t = \text{Reflect}(S_t, A_t, T_t, \delta, \epsilon) := \arg\max_H \left[P\left(S_{t+k}^{\text{obs}} \mid H\right) - P\left(S_{t+k}^{\text{pred}} \mid C\right) \right] \tag{7}$$

Here, H denotes a candidate hypothesis, such as a shift in a causal dependency, an unmodeled confounder, or a misestimated perturbation factor δ. These hypotheses are interpretable and testable, grounded in structured reasoning rather than neural approximation.

To close the loop between formal inference and explainability, we introduce a modular architecture for Reflective Causal Agents, composed of three distinct modules:

- **Causal Inference Engine** models dynamic systems with the function

  $$C(S_t, A_t, T_t, \delta) \to S_{t+k},$$

  capturing delayed and nonlinear effects
- **Reflect Mechanism** performs structured self-correction by generating and evaluating causal hypotheses when discrepancies are detected
- **LLM-Based Interpreter** translates formal causal tuples and hypotheses into natural language explanations or counterfactuals, acting as a generative interface with users.

This architecture explicitly separates inference from interpretation. The causal engine governs reasoning; the language model merely reports its conclusions. This prevents LLMs from hallucinating ungrounded causal chains, a known limitation in language-based agents. Instead, explanations take the form:

> "Action: A_t was expected to lead to S_{t+k}, but instead resulted in S'_{t+k}.
> Hypothesis: perturbation δ shifted the influence of A_t, activating an alternate causal pathway."

By design, Reflective Causal Agents do not simply optimize for external reward. They aim to build and revise an internal model of the environment's causal dynamics. This shift from policy optimization to causal understanding positions them for stronger generalization, interpretability, and alignment especially in non-stationary environments where brittle policies and black-box reasoning often fail.

5 Comparison with Existing Approaches

To situate our framework, we compare it with traditional reinforcement learning and contemporary causal reinforcement learning approaches. Causal Reflection targets dynamic, time-varying causal relationships. Unlike typical Casual Reinforcement Learning (CRL) that uses static graphs, our temporal causal function models non-stationary environments. Including time T and perturbation

Algorithm 1 Reflect Mechanism for Causal Hypothesis Generation

Require: Current state S_t, action A_t, time T_t, perturbation δ, causal function C, observed outcome S_{t+k}^{obs}, loss function L, threshold τ
Ensure: Set of causal hypotheses H_t

$\hat{S}_{t+k} \leftarrow C(S_t, A_t, T_t, \delta)$
$\epsilon \leftarrow L(\hat{S}_{t+k}, S_{t+k}^{obs})$
if $\epsilon > \tau$ **then**
  Generate candidate hypotheses $\mathcal{H}$ to explain discrepancy
  $H_t \leftarrow \arg\max_{H \in \mathcal{H}} \big[P(S_{t+k}^{\text{obs}} \mid H) - P(\hat{S}_{t+k} \mid C)\big]$
  for each $H \in H_t$ **do**
    Test hypothesis H
    if H is valid **then**
      Update causal function C with H
    end if
  end for
else
  $H_t \leftarrow \emptyset$
end if
return H_t

δ helps capture evolving causal dynamics which is a key challenge in CRL research. Our Causal Reflection framework introduces fundamental shifts in the objective, mechanism, and capabilities of an intelligent agent.(see Table 1).

The table below summarizes these differences.

Table 1: Comparison of Decision-Making Paradigms

Dimension	Traditional RL (PPO)	Causal RL (CRL)	Causal Reflection
Primary Goal	Maximize cumulative reward.	Improve policy learning (sample efficiency, generalization) using a causal model.	Build an accurate, dynamic causal model for explanation and prediction.
Core Mechanism	Policy optimization via trial-and-error.	Causal discovery/ inference on a static world model to inform policy.	Predictive modeling and causal hypothesis testing on a dynamic world model.
Handling of Time	Sequential states, but no explicit model of temporal causality.	Often assumes a static, time-invariant causal graph.	Explicitly models temporal delays and time-varying dynamics.
Role of LLM	N/A, or used for auxiliary tasks like reward shaping.	Can be a source of prior knowledge for the static causal graph.	A generative inference engine that translates the formal causal model's output into natural language explanations.

6 Applications and Use Cases

The Causal Reflection framework offers practical benefits for building intelligent, explainable, and aligned AI systems, demonstrated in two main applications.

Causal Self-Reflection for Explainability: Our framework enables deep causal explanations beyond simple correlations. For example, a user logging states like "burnout," actions like "taking on extra

projects," and timestamps can receive explanations generated by the LLM. Guided by the temporal causal model, the LLM interprets the data to produce structured hypotheses explaining burnout as a cumulative effect influenced by hidden factors (δ) such as stress or poor sleep. This interaction yields actionable, formal causal insights.

Human-AI Alignment via Causal Hypotheses: Rather than just matching goals, the LLM uses the causal model to reason about underlying mechanisms. For a user wanting to boost productivity, it generates causal hypotheses based on past (S, A, T, δ) data like unplanned tasks causing overwhelm after 24 hours or poor sleep intensifying workload effects and articulates targeted interventions. This shifts AI support from surface-level advice to meaningful causal understanding and personalized guidance.

7 Discussion and Future Work

While our framework offers a promising new direction, we believe that the theoretical contribution that opens several avenues for future research. This section discusses a proposed validation strategy, acknowledges open challenges, and outlines future research directions.

Learning the Causal Function: This paper has posited the theoretical existence and form of the causal function C. A critical next step is to develop methods for learning its parameters from a combination of observational and intervention data.

Validation Strategy: We propose the development of a benchmark suite of simulated environments with known, dynamically changing causal relationships to validate our framework. For instance, a simulated economic system could model government intervention as a perturbation (δ) that temporarily alters the causal link between interest rates (Action) and inflation (State). Evaluation of such an agent implementing Causal Reflection focuses not on traditional reward metrics, but on its ability to:

1. **Identify the active causal graph:** How quickly and accurately does the agent detect and update its causal model following a structural break?
2. **Predict future states:** How effectively does the agent minimize prediction error, particularly after perturbations?

Extension to Multi-Agent Systems: A compelling future direction is to extend the Causal Reflection framework to Multi-Agent Reinforcement Learning (MARL). In such settings, an agent must model not only the causal impact of its own actions but also the causal influence of other agents' actions on the environment and on each other. Applying causal reasoning to MARL [Briglia et al., 2025] is a nascent but critical research area for understanding complex social and strategic interactions.

We treat this paper as a formal foundation for a forthcoming implementation, Causal Reflection Agents, a simulation suite to evaluate their performance against LLMs and RL Agents.

8 Conclusion

We introduced *Causal Reflection*, a framework that shifts focus from reward maximization to building accurate, interpretable causal models of dynamic environments. By modeling state, action, time, and perturbations, our approach captures nonlinear, time-varying causal relationships. We also outlined how LLMs can serve as generative engines to translate these formal outputs into structured, natural language explanations. This framework lays the groundwork for more robust, adaptive, and explainable AI systems aligned with human reasoning.

9 Limitations

Several challenges must be addressed to fully realize the potential of this framework, two of the most urgent and critical ones being-

Scalability: Modeling complex, high-dimensional systems is computationally intensive. The state representation S_t can become prohibitively large, and inferring the causal function C in such spaces is a significant challenge, echoing broader issues in high-dimensional causal inference. Future work

should explore factorization and representation learning techniques to create lower-dimensional, causally sufficient state spaces. While reasoning over dynamic causal models increases complexity, the framework is modular: δ estimation and *Reflect* are invoked conditionally (only when prediction error is high), and causal model updates are localized. For high-dimensional state spaces, we recommend applying dimensionality reduction techniques (e.g., Variational Autoencoders (VAEs) or causal autoencoders) to obtain compact representations.

LLM Fidelity and Controllability: The framework relies on the LLM to be a faithful interpreter of the formal model's output. However, LLMs can "hallucinate" or misrepresent information. Research is needed to develop methods for quantifying and mitigating these "translation errors" to ensure the natural language explanations remain rigorously grounded in the underlying causal inference.

As more experiments and adaptions of Causal Reflection roll in, we are likely to see more limitations discussed.

References

Emre Kiciman, Robert Ness, Amit Sharma, and Chenhao Tan. Causal reasoning and large language models: Opening a new frontier for causality. *Transactions on Machine Learning Research*, 2023.

Maximilian Seitzer, Bernhard Schölkopf, and Georg Martius. Causal influence detection for improving efficiency in reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 34, pages 22905–22918, 2021.

Lijing Jiao, Ya-nan Wang, Xuan Liu, Lisha Li, Fang Liu, Wenping Ma, and Boyu Hou. Causal inference meets deep learning: A comprehensive survey. *Research*, 7:0467, 2024.

Siqi Du, Guang Song, Lixin Han, and Han Hong. Temporal causal inference with time lag. *Neural computation*, 30(1):271–291, 2017.

Zheyuan Deng, Jun Jiang, Guodong Long, and Chengqi Zhang. Causal reinforcement learning: A survey. *arXiv preprint arXiv:2307.01452*, 2023.

Geyu Wan, Yixin Lu, Yewen Wu, Min Hu, and Si Li. Large language models for causal discovery: Current landscape and future directions. *arXiv preprint arXiv:2402.11068*, 2024.

Xiao Liu, Peng Xu, Junda Wu, Jinyang Yuan, Yanchi Yang, Yi Zhou, and Furong Huang. Large language models and causal inference in collaboration: A comprehensive survey. In *Findings of the Association for Computational Linguistics: NAACL 2025*, pages 7668–7684, 2025.

Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of Causal Inference: Foundations and Learning Algorithms*, chapter 9. The MIT Press, 2017.

Zhiming He, Song Li, Wen Song, Le Yuan, Jie Liang, Hong Li, and Keke Gai. Learning time-aware causal representation for model generalization in evolving domains. *arXiv preprint arXiv:2506.17718*, 2025.

Elias Bareinboim, Jiji Zhang, and Sang-Hyeun Lee. An introduction to causal reinforcement learning. *arXiv preprint arXiv:2101.06498*, 2021.

Alberto Méndez-Molina, Efrén F. Morales, and L. Enrique Sucar. Causal discovery and reinforcement learning: A synergistic integration. In *International Conference on Probabilistic Graphical Models*, pages 421–432. PMLR, 2022.

Nikolaos Gkorgkolis, Nikolaos Kougioulis, Minlan Wang, Baris Caglayan, Alberto Tonon, Diego Simionato, and Ioannis Tsamardinos. Temporal causal-based simulation for realistic time-series generation. *arXiv preprint arXiv:2506.02084*, 2025.

Jonathan Calderon and Gordon J. Berman. Inferring the time-varying coupling of dynamical systems with temporal convolutional autoencoders. *arXiv preprint arXiv:2406.03212*, 2024.

Milan Paluš, Anna Krakovská, Jaroslav Jakubík, and Martina Chvosteková. Causality, dynamical systems and the arrow of time. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(7), 2018.

Jason Gkountouras, Moritz Lindemann, Phillip Lippe, Efstratios Gavves, and Ivan Titov. Language agents meet causality–bridging llms and causal world models. *arXiv preprint arXiv:2410.19923*, 2024.

Linyi Wang and Yuxi Shen. Evaluating causal reasoning capabilities of large language models: A systematic analysis across three scenarios. *Electronics*, 13(23):4584, 2024.

Shubham Ashwani, Kavya Hegde, Narendra Reddy Mannuru, Divyansh Singh Sengar, Mayank Jindal, K. Chaitanya Reddy Kathala, and Aseem Chadha. Cause and effect: can large language models truly understand causality? In *Proceedings of the AAAI Symposium Series*, volume 4, pages 2–9, 2024.

Heng Chi, Haotian Li, Wen Yang, Feiyi Liu, Lihui Lan, Xiang Ren, and B. Han. Unveiling causal reasoning in large language models: Reality or mirage? In *Advances in Neural Information Processing Systems*, volume 37, pages 96640–96670, 2024.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 36, pages 8634–8652, 2023.

Giovanni Briglia, Stefano Mariani, and Franco Zambonelli. A roadmap towards improving multi-agent reinforcement learning with causal discovery and inference. *arXiv preprint arXiv:2503.17803*, 2025.