Hypothesis Generation and Intervention Design for Smart Environments with Causal Reasoning in Large Language Models

Taqiya Ehsan, Jorge Ortiz
Department of Electrical & Computer Engineering, Rutgers University
taqiya.ehsan@rutgers.edu, jorge.ortiz@rutgers.edu

***Abstract*—This paper investigates whether integrating causal knowledge into Large Language Models (LLMs) enhances their ability to generate hypotheses and design effective interventions in smart environments. Specifically, we compare a causal agent, equipped with Directed Acyclic Graphs (DAGs) and counterfactual reasoning, against a baseline agent without explicit causal knowledge. Using a smart room simulation, we evaluate their performance across metrics such as relevance, novelty, alignment, and efficiency. Our results demonstrate that incorporating causal reasoning improves the accuracy and robustness of LLMs in optimizing energy efficiency and user satisfaction, highlighting the potential of causal-aware AI systems for complex decision-making environments.**

I. Introduction

Smart environments, such as homes and offices, rely on dynamically managing variables like temperature, humidity, and air quality to optimize energy efficiency and user satisfaction. Traditional control systems often struggle to balance these factors, especially in scenarios with complex and dynamic causal relationships.

Large Language Models (LLMs) have shown promise in adaptive reasoning tasks, including hypothesis generation and decision-making. However, their ability to effectively navigate complex systems remains limited without explicit access to causal knowledge. This research investigates whether equipping LLMs with causal reasoning, represented through Directed Acyclic Graphs (DAGs) and counterfactual analysis, enhances their performance in hypothesis generation and intervention design for smart environments.

We compare two agents: a *Causal Agent* with access to causal relationships and reasoning capabilities, and a *Baseline Agent* that relies only on pattern-based inference. The agents are evaluated in a simulated smart room environment on metrics such as relevance, novelty, and alignment of hypotheses and interventions. This work contributes to the growing field of causal-aware AI by exploring the integration of causal inference with LLMs in practical applications.

II. Problem Statement

Managing dynamic variables such as temperature (T), humidity (H), and air quality (AQ) in smart environments is critical to achieving two primary outcomes:

- **Energy Efficiency** (E): Minimizing energy consumption.
- **User Satisfaction** (S): Maximizing user comfort.

Traditional control systems struggle with balancing these factors, particularly in dynamic settings with interdependent causal relationships. For example, reducing energy consumption may inadvertently lower user satisfaction by causing discomfort, while improving air quality may increase energy costs.

This research addresses whether providing causal knowledge to LLMs can improve:

1) Hypothesis generation for identifying key causal relationships.
2) Intervention design for optimizing outcomes while adhering to system constraints.

Two LLM-based agents are evaluated:

- **Causal Agent** (A_C): Equipped with Directed Acyclic Graphs (DAGs) and counterfactual reasoning.
- **Baseline Agent** (A_B): Relies on heuristic-based, pattern-matching approaches without explicit causal information.

III. Mathematical Framework

A. Key Components

- **State Variables** ($\mathbf{X}$): Environmental conditions in the smart room:
$$\mathbf{X} = \{T, H, AQ\},$$
where T, H, and AQ represent temperature, humidity, and air quality, respectively.
- **Control Variables** ($\mathbf{I}$): Interventions to modify the environment:
$$\mathbf{I} = \{I_T, I_H, I_{AQ}\}.$$
- **Outcome Variables**: Metrics of interest:
  - Energy Consumption ($E(\mathbf{X})$)
  - User Satisfaction ($S(\mathbf{X})$)
- **Causal Relationships** (G): A Directed Acyclic Graph (DAG) represents dependencies:
$$\mathbf{X} \to \{E, S\}.$$
- **Constraints**: The environmental variables are bounded by:
$$T \in [18, 30]^\circ C, \quad H \in [30, 70]\%, \quad AQ \in [0, 500].$$

B. Reward Function

The reward for an intervention is a weighted sum of user satisfaction (S) and energy efficiency (E):

$$R(\mathbf{I}) = \alpha S(\mathbf{X}) - \beta E(\mathbf{X}),$$

where $\alpha > 0$ and $\beta > 0$ are weights prioritizing satisfaction and efficiency, respectively.

C. Risk Function

Risk corresponds to deviations from ideal operating conditions:

$$\text{Risk}(\mathbf{I}) = \sum_i \lambda_i \cdot \left(X_i - X_i^{\text{ideal}}\right)^2,$$

where $\lambda_i > 0$ penalizes deviations from ideal values X_i^{ideal}.

D. Optimization Problem

The objective is to maximize reward while minimizing risk:

$$\max_{\mathbf{I}} \; \mathbb{E}[R(\mathbf{I})] - \gamma \cdot \text{Risk}(\mathbf{I}),$$

where $\gamma > 0$ balances the trade-off between reward and risk.

E. Causal Agent vs. Baseline Agent

- **Causal Agent** (A_C): Utilizes the DAG (G) to estimate:

  $$\mathbb{E}[S(\mathbf{X})] \quad \text{and} \quad \mathbb{E}[E(\mathbf{X})],$$

  incorporating counterfactual reasoning to refine interventions.
- **Baseline Agent** (A_B): Relies on heuristic-based predictions without access to explicit causal information.

IV. Related Work

A. Integration of Causal Inference and Large Language Models

Recent research has explored the intersection of causal inference and LLMs, aiming to enhance the reasoning capabilities of AI systems. For example:

1) **LLM-Guided Causal Discovery**: Vashishtha et al. [1] demonstrated how LLMs can estimate causal effects using virtual expert methods for causal ordering of variables.
2) **Causal Inference Benchmarking**: Jin et al. [2] introduced CLadder, a benchmark for assessing causal reasoning in LLMs across associational, interventional, and counterfactual tasks.

B. Causal Reasoning Capabilities of LLMs

Other works focus on improving the causal reasoning capacity of LLMs:

1) Xu et al. [3] surveyed the integration of causal inference with LLMs, emphasizing multimodal reasoning and fairness.
2) Ma [4] evaluated LLMs on diverse causal tasks, including pairwise discovery and counterfactual reasoning.
3) Kıcıman et al. [5] explored LLMs' capabilities in generating correct causal arguments with high probability.

C. Applications in Complex Environments

The integration of causal reasoning in LLMs is particularly promising for applications like smart environments. Xu et al. [3] highlighted the relevance of multimodal data sources, while other works propose combining statistical and knowledge-based causal inference techniques [6], [7].

V. Methodology

A. Problem Formulation

This study addresses two primary outcomes in a smart room environment:

- **Energy Efficiency** (E): Minimize energy consumption.
- **User Satisfaction** (S): Maximize user comfort.

Key influencing variables include temperature (T), humidity (H), and air quality (AQ).

B. Agents

- **Causal Agent** (A_C): Leverages DAGs and counterfactual reasoning for hypothesis generation and intervention design.
- **Baseline Agent** (A_B): Relies on pattern recognition without access to causal information.

C. Simulation Environment

A Node.js-based simulator replicates dynamic environmental conditions. Key features include:

- Real-time updates with millisecond precision.
- Variable bounds: $T \in [18, 30]°C$, $H \in [30\%, 70\%]$, $AQ \in [0, 500]$.
- Immediate feedback on interventions.

D. Workflow

1) Agents generate hypotheses (H) about causal relationships.
2) Propose interventions (I) to modify environmental variables.
3) Evaluate outcomes based on predefined metrics.

VI. Evaluation Metrics

A. Hypothesis Evaluation Metrics

- **Relevance Score:** Measures alignment with ground truth DAGs.
- **Novelty Score:** Quantifies unique insights introduced in hypotheses.
- **Breadth Score:** Evaluates the comprehensiveness of covered causal paths.

B. Intervention Evaluation Metrics

- **Impact Accuracy:** Assesses the alignment between predicted and observed outcomes.
- **Efficiency Score:** Measures proximity to ideal values for key metrics.
- **Alignment Score:** Combines coordination, moderation, and efficiency metrics into a single score.

VII. Results and Discussion

The causal agent (A_C) consistently outperformed the baseline agent (A_B) across all metrics:

- **Relevance and Breadth:** A_C achieved 25% higher relevance scores and demonstrated broader causal reasoning.
- **Impact Accuracy:** A_C's intervention predictions were 30% more accurate.
- **Novelty:** A_C introduced 40% more novel elements in its hypotheses.

A. Analysis

The superior performance of A_C highlights the value of incorporating causal knowledge into LLMs. The ability to reason about both direct and indirect relationships enabled more precise and impactful interventions.

VIII. Conclusion and Future Work

This work demonstrates the efficacy of causal reasoning in enhancing LLMs' ability to generate hypotheses and interventions for dynamic environments. Future work will explore:

- Integration with multimodal data (e.g., visual and sensory inputs).
- Real-time adaptive learning for evolving environments.

References

[1] S. Vashishtha, S. Yadlowsky, and S. Athey, "Estimating causal effects with large language models as virtual experts," *arXiv preprint arXiv:2310.09851*, 2023.

[2] Z. Jin, K. Singhal, T. Fog, N. Garg, L. Xiao, S. Subramanian, D. Sontag, A. Tamkin, B. Schölkopf, B. Barak *et al.*, "Cladder: Assessing causal reasoning in language models," *arXiv preprint arXiv:2312.04350*, 2023.

[3] H. Xu *et al.*, "A survey on large language models and causal inference: Towards causal ai," *arXiv preprint arXiv:2401.03721*, 2024.

[4] Y. Ma, "Large language models for causal inference," *arXiv preprint arXiv:2402.07962*, 2024.

[5] E. Kıcıman, R. Ness, A. Sharma, and C. Tan, "Causal reasoning and large language models: Opening a new frontier for causality," *arXiv preprint arXiv:2305.00050*, 2023.

[6] J. Takayama and S. Shimizu, "Statistical causal prompting: Integrating statistical causal discovery with knowledge-based causal inference using large language models," *arXiv preprint arXiv:2402.01173*, 2024.

[7] Y. Zhang, Y. Zhu, Y. Zhu, X. Wang, Y. Guo, S. Gao, and H. Xu, "Utilizing llms for enhanced argumentation and extraction of causal knowledge in disease mechanism studies," *medRxiv*, 2024.