

SCOOP: A Framework for Proactive Collaboration and Social Continual Learning through Natural Language Interaction and Causal Reasoning

Dimitri Ognibene, Sabrina Patania, Luca Annese, Cansu Koyuturk, Franca Garzotto, Giuseppe Vizzari
 University of Milan-Bicocca
 Milan, Italy
 dimitri.ognibene@unimib.it

Azzurra Ruggeri
 TUM School of Social Sciences and Technology
 Munich, Germany
 a.ruggeri@tum.de

Simone Colombani
 Oversonic Robotics
 Carate Brianza, Italy

Abstract

Multimodal information-gathering settings, where users collaborate with AI in dynamic environments, are increasingly common. These scenarios involve complex processes with textual and multimodal interaction (e.g., house refurbishment plans) and often require accessing additional structural information (e.g., regulations) via cost-incurring requests. Moreover, AI helpers lack access to users' true goals, beliefs, and preferences and struggle to integrate diverse information effectively.

We propose a social continual learning framework for causal knowledge acquisition and collaborative decision-making. It focuses on autonomous agents learning through dialogues, question-asking, and interaction in open, partially observable environments. A key component is a natural language oracle that answers the agent's queries about environmental mechanisms and states, refining causal understanding while balancing exploration (learning) and exploitation (using knowledge).

Evaluation tasks, inspired by developmental psychology, emphasize causal reasoning and question-asking skills, complementing benchmarks by assessing the agent's ability to identify knowledge gaps, generate meaningful queries, and incrementally update reasoning. The framework also evaluates how the cost of acquiring knowledge is amortized across tasks in the same environment.

We propose two architectures: (1) a system combining Large Language Models (LLMs) with the ReAct framework and question-generation, and (2) an advanced system with a causal world model (symbolic, graph-based, or subsymbolic) for reasoning and decision-making. The latter builds a causal knowledge graph for efficient inference and adaptability under constraints. Challenges include integrating causal reasoning into ReAct and optimizing exploration and question-asking in error-prone scenarios. Beyond applications, this framework models developmental processes combining causal reasoning, question generation, and social learning.

Keywords: social learning, question generation, collaborative AI, LLM, planning, continual learning, cognitive development.

Acknowledgements

We acknowledge support from the Volkswagen Foundation for the project *Developing an Artificial Social Childhood (ASC) to improve AI causal reasoning, information gathering and decision making*, Ref.: 9E530.

1 Introduction

With the advent of advanced generative AI, tasks involving rich multimodal and natural language information—where users collaborate with AI helpers—are becoming increasingly common. These settings often involve:

1. **Sequential Collaborative Interactions:** performing the task requires planning and executing of sequences of interactions and independent by the user or the agent (e.g. collecting different elements and tools), with the need to account for their consequences over multiple steps.
2. **Costly Social Information Access:** users and agents must manage limited resources when accessing information provided by other agents (e.g., mechanisms descriptions, regulations or user preferences or beliefs [1]).
3. **Multimodal complexity:** combining textual descriptions, visual representations, and structural information (e.g., house refurbishment plans).
4. **Dynamic Environments:** evolving tasks with shifting states and requirements (e.g., bureaucratic processes, repair scenarios) [27].

The simultaneous presence of these conditions and their interaction motivate our work to formalize and create effective solutions. Moreover, they reflect important aspects of human developmental processes [3, 6].

1.1 Example Scenarios

Example 1: Repairing and Shipping Items

A robot agent assists in repairing and shipping items to various international destinations. Tasks include adhering to specific repair requirements, customs regulations, packaging standards, and climate conditions. The agent collaborates with the user to clarify repair strategies and queries a natural language oracle for regulatory or technical information. During downtime, the agent explores tools and equipment to refine its understanding of repair techniques, improving future performance.

Example 2: Compiling Documents for Official Requests

An agent aids a user in compiling documents for official applications, such as visas or tax submissions. Tasks involve identifying required documents, extracting relevant information, and validating regulations via a natural language oracle. When ambiguities arise, the agent interacts with the user for clarification. In idle periods, it autonomously explores document templates and regulatory archives to improve performance across multiple problem instances.

2 Related Work

ReAct Framework: The ReAct (Reason + Act) framework [26] integrates reasoning and acting capabilities into LLMs, enabling contextual decision-making and blending question generation with action-taking in collaborative environments.

Causal Reasoning in AI: Techniques like causal discovery [12] and inference frameworks such as DoWhy [32] provide essential tools for reasoning in dynamic, partially observable environments.

Integrating Causal Graphs with LLMs: Recent research integrates causal graphs with LLMs to enhance reasoning and decision-making. Embedding causal reasoning into LLM workflows enables systems to interpret environments, predict outcomes, and optimize actions. Frameworks combining LLMs with causal world models [11] and studies on automating causal discovery [23] highlight the potential of causal representation learning for dynamically constructing world models.

LLMs and Planning Systems: Combining LLMs with planning systems, such as symbolic planners or reinforcement learning frameworks, supports structured decision-making. Hybrid approaches integrate LLMs for language understanding and planners for task execution, excelling in complex scenarios like multi-agent collaboration and robotics [22, 30].

Causal Discovery and Question Generation: Integrating LLMs with causal discovery tools enhances reasoning in dynamic environments. For instance, PyWhy-LLM supports causal analysis [9], while DoWhy-GCM facilitates inference in graphical causal models [19]. These tools enable LLMs to refine causal graphs and resolve ambiguities, closing the loop between knowledge acquisition and action.

Symbolic and Subsymbolic Integration: Hybrid architectures bridge symbolic causal reasoning with the subsymbolic capabilities of LLMs, combining graph-based reasoning with unstructured data processing for robust decision-making agents [20, 16].

Active Information Gathering: Research on active learning and information gain [24, 15, 14, 8, 10] informs strategies for querying to maximize utility while balancing exploration and exploitation in interactive systems.

Human-in-the-Loop Systems: Human feedback integration highlights the value of interactive learning, enabling agents to dynamically query users and adapt to preferences [21].

Developmental Psychology Insights: Insights from children’s causal learning [5, 31] inspire benchmarks to evaluate AI agents’ abilities in question generation and causal inference.

By synthesizing these elements, our framework advances adaptive, interactive AI systems at the intersection of causal reasoning, social learning, and multimodal interaction.

3 The SCOOP framework: Social Continual Object-Oriented POMDP

Formal Framework. We formalize our rich interactive setting as an *object-oriented* partially observable Markov decision process (OO-POMDP), extended to incorporate a *linguistic world descriptor* generating natural language observations (e.g. the initial task description in natural language), multiple problem instances, a *user* with problem-specific objectives, and a *natural language oracle* that provides causal information. Let

$$\mathcal{D} = (\mathcal{T}_O, \mathcal{F}_O, \mathcal{R}_O, \mathcal{W}_O)$$

denote a *domain specification*, where:

- \mathcal{T}_O is a set of object types (e.g., boxes, tools, containers);
- \mathcal{F}_O is the set of allowed features or predicates relevant to the domain (e.g., “contains(x,y)” or “isOpenable(x)’);
- \mathcal{R}_O is an evolving set of *causal rules* capturing how these features and object types interact (e.g., whether opening a container allows access to its contents), some of which may be unknown or partially specified;
- \mathcal{W}_O is the family of possible world configurations consistent with $\mathcal{T}_O, \mathcal{F}_O$, and \mathcal{R}_O .

A *problem instance* θ refines \mathcal{D} with a concrete set of objects, an initial state, and a *user objective* (defined by a problem-specific reward function r_θ that encodes the user’s goals). The *helper agent* interacts with both the user, making questions to clarify his goals or preferences, presenting results and suggestions, or other interactive actions, and a *natural language oracle* (to acquire missing causal rules or environment states) across potentially many problem instances $\{\theta_i\}$. The oracle responds in multiple formats:

1. *Language-based descriptions* of environment dynamics, such as “box A must be opened before retrieving item B”;
2. *Formal causal chunks*, where the oracle may directly provide rules or graphs parts (e.g., “node Open(Box) causes Accessible(Item)”),
3. *Observation-like feedback*, akin to sensor readings or state confirmations.

The overall action space $\mathcal{A} = \mathcal{A}^a \cup \mathcal{A}^u$ composed by the agent’s *action space* by \mathcal{A}^a and the user’s *action space* by \mathcal{A}^u . We denote the agent’s *action space* by $\mathcal{A}^a = \mathcal{A}_{act}^a \cup \mathcal{A}_{query}^a$, where \mathcal{A}_{act} includes environment-oriented actions (e.g., open, pick, place) and \mathcal{A}_{query} includes queries to the oracle or the user. The agent also observes the user’s feedback or clarifications regarding the task objectives and environment states. We denote the user’s *action space* by $\mathcal{A}^u = \mathcal{A}_{act}^u \cup \mathcal{A}_{query}^u$, where \mathcal{A}_{act}^u includes environment-oriented actions (e.g., open, pick, place) and \mathcal{A}_{query}^u includes queries to the agent. Formally, each problem instance is modeled as:

$$(\mathcal{S}_\theta, \mathcal{A}, \Omega_\theta, \mathcal{T}_\theta, \mathcal{O}_\theta, r_\theta^u, r_\theta^a, \gamma, \beta),$$

mirroring a POMDP with the following modifications:

- \mathcal{S}_θ embeds the object-based states from θ and any partial knowledge of the causal rules \mathcal{R}_O ;
- Ω_θ is the space of possible observations, spanning both *environmental signals* (e.g., sensor readings, gripper state, environment map chunk, etc) and *language-based responses* from the user or oracle;
- r_θ^u encodes the *user’s objectives* for instance θ , which the helper agent aims to optimize;
- r_θ^a encodes the *helper agent’s action costs* for instance θ , which the helper agent aims to optimize;
- $\beta(\cdot)$ defines the *cost of querying* (time, resources, or complexity) the oracle to obtain new causal information or the user about current objective and preferences, i.e. r_θ^u .

Crucially, the agent can *explore* the domain outside active tasks to refine \mathcal{R}_O (e.g., by performing experiments or asking domain-level questions). Any information gleaned is *amortized* across future tasks θ_j . This design enables **continual learning** of domain mechanics: as the agent accumulates causal knowledge (e.g., “a certain box can contain items of type

T''), it improves performance in subsequent problem instances. More formally this is obtained assuming that at each instance θ_j the specific instantiation of \mathcal{R}_O , r_θ^u and r_θ^a are extracted from the same distribution.

Ultimately, the helper agent objective functions is: $\sum_\theta \sum_t^{T(\theta)} \gamma^{t+T(-\theta)} (r_\theta^u(s_t, a_t^u) + r_\theta^a(s_t, a_t^a) + \beta(a_t^a))$. Balancing *exploration* (question-asking, active experimentation) and *exploitation* (leveraging current knowledge to solve tasks efficiently) is thus a central challenge in this social continual learning framework.

4 Developmental Psychology-Inspired Tasks for Evaluating Causal Reasoning and Question-Making

Drawing from developmental psychology, we design tasks to evaluate causal reasoning and question-making skills in collaborative AI systems. Inspired by children’s learning behaviors, these tasks assess the agent’s ability to:

Explore-Exploit Tradeoff: Balance between directed exploration and utilizing known information to reduce uncertainty and achieve goals efficiently [7]. *“Why” and “What If” Questions*: Formulate meaningful hypotheses and evaluate counterfactual scenarios to refine causal models [33]. *Epistemic Question Formulation*: Construct precise, goal-directed queries to address knowledge gaps efficiently [25]. *Causal Inference and Learning*: Engage AI in tasks where it observes incomplete sequences of events and must infer causal relationships. For example, after observing that certain components drive a machine, the AI predicts outcomes without direct trial-and-error [4, 18]. *Generating Hypotheses from Confounded Evidence*: Assesses AI’s ability to generate interventions that are informative in resolving the structure of ambiguous causal system [29]. These tasks provide a multidimensional framework to benchmark AI systems, focusing on cognitive, linguistic, and social reasoning capabilities essential for dynamic, real-world collaboration.

5 Reference architectures

5.1 Base: Oracle-Aided ReAct

The base architecture simply extends the ReAct framework introducing actions to ask state and the user about their preference and objectives and the oracle about environment mechanics. However, as noted in the literature, complex planning [17] and usage of declarative knowledge for decision making and execution [2] appear difficult for vanilla LLMs.

5.2 Advanced: ReAct Framework with Oracle-Aided Causal Reasoning

Building on the base architecture, the advanced ReAct (Reason + Act) [26] framework introduces information-gathering actions and extends its functionality with a specialized action, `CausalRefinementAndAction`. This action is invoked by the Large Language Model (LLM) when complex reasoning tasks are required, specifically for:

- **Refining or updating** knowledge about user needs and the world’s mechanisms and states (causal model), or
- **Planning and executing** steps to achieve a specified goal.

`CausalRefinementAndAction` integrates iterative causal knowledge management, utilizing both external oracle support (e.g., a domain expert or automated simulator) and established causal inference libraries such as *causal-learn*, *DoWhy*, and *Tetrad* [28, 32, 13]. Given a user’s prompt, current goal, and contextual information, the LLM initially maps relevant knowledge to a causal graph, which may be incomplete.

The agent estimates the expected value and cost of potential actions to refine its knowledge, using metrics such as Value of Information (VoI) or robust optimization criteria. Refinement actions include querying the user about preferences and goals, asking the oracle about specific causal links or effect sizes, or performing interventions. If the refinement is deemed beneficial (i.e., cost is below a threshold), the suggested strategy is executed. Based on responses from the user or oracle, or results of interventions, causal inference libraries update the graph and determine whether additional refinements are necessary.

Once the causal graph is sufficiently refined, the ReAct agent invokes planning routines—using libraries such as *PyCID* or a robust Markov Decision Process (MDP) solver—to derive policies or action sequences that maximize the likelihood of achieving the user’s goals under the current causal knowledge. This combination of LLM-driven reasoning, causal knowledge management, and decision-making enable advanced reasoning and information-gathering capabilities are activated when necessary while maintaining the flexibility to handle diverse scenarios typical of LLMs.

References

- [1] F. Bianco and D. Ognibene. From psychological intention recognition theories to adaptive theory of mind for robots: Computational models. In *HRI*, pages 136–138, 2020.
- [2] A. Arabi et al. Habit coach: Customising rag-based chatbots to support behavior change. *arXiv*, 2024.
- [3] A. Cangelosi et al. Integration of action and language knowledge: A roadmap for developmental robotics. *IEEE Transactions on Autonomous Mental Development*, 2(3):167–195, 2010.
- [4] A. Gopnik et al. Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Dev. Psychol.*, 37(5):620, 2001.
- [5] A. Gopnik et al. A theory of causal learning in children: causal maps and bayes nets. *Psychol. Rev.*, 111(1):3, 2004.
- [6] A. Ruggeri et al. Sources of developmental change in the efficiency of information search. *Developmental psychology*, 52(12):2159, 2016.
- [7] B. Meder et al. Development of directed and random exploration in children. *Dev. Sci.*, 24(4):e13095, 2021.
- [8] D. Ognibene et al. Proactive intention recognition for joint human-robot search and rescue missions through monte-carlo planning in pomdp environments. In *Social Robotics: 11th International Conference, ICSR*, pages 332–343. Springer, 2019.
- [9] E. Kiciman et al. Causal reasoning and large language models: Opening a new frontier for causality. *arXiv*, 2023.
- [10] E. Masiero et al. In search of compositional multi-task deep architectures for information theoretic field exploration. In *2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)*, pages 612–617. IEEE, 2024.
- [11] J. Gkountouras et al. Language agents meet causality – bridging llms and causal world models, 2024.
- [12] J. Peters et al. *Elements of causal inference: foundations and learning algorithms*. MIT Press, 2017.
- [13] J. Ramsey et al. Tetrads—a toolbox for causal discovery. In *Climate Informatics*, pages 1–4, 2018.
- [14] K. Friston et al. Active inference and epistemic value. *Cogn. Neurosci.*, 6(4):187–214, 2015.
- [15] L. Bertolazzi et al. Chatgpt’s information seeking strategy: Insights from the 20-questions game. In *INLG*, pages 153–162, 2023.
- [16] M. Ibrahim et al. Special session: Neuro-symbolic architecture meets large language models: A memory-centric perspective. In *CODES+ISSS*, pages 11–20, 2024.
- [17] M. Katz et al. Thought of search: Planning with language models through the lens of efficiency. In *NeurIPS*, 2024.
- [18] M. Shavlik et al. Contributions of causal reasoning to early scientific literacy. *J. Exp. Child Psychol.*, 224:105509, 2022.
- [19] P. Blöbaum et al. Dowhy-gcm: An extension of dowhy for causal inference in graphical causal models. *JMLR*, 25(147):1–7, 2024.
- [20] P. Hitzler et al. Neuro-symbolic approaches in artificial intelligence. *Natl. Sci. Rev.*, 9(6):nwac035, 2022.
- [21] S. Amershi et al. Power to the people: The role of humans in interactive machine learning. *AI Mag.*, 35(4):105–120, 2014.
- [22] S. Colombani et al. One to rule them all: natural language to bind communication, perception and action. *arXiv*, 2024.
- [23] S. Long et al. Can large language models build causal graphs?, 2024.
- [24] S. Patania et al. Large language models as an active bayesian filter: information acquisition and integration. In *SemDial*, 2024.
- [25] S. Ronfard et al. Question-asking in childhood: A review of the literature and a framework for understanding its development. *Dev. Rev.*, 49:21–39, 2018.
- [26] S. Yao et al. React: Synergizing reasoning and acting in language models, 2023.
- [27] T. Taniguchi et al. World models and predictive coding for cognitive and developmental robotics: Frontiers and challenges. *Adv. Robot.*, 37(13):780–806, 2023.
- [28] Y. Zheng et al. Causal-learn: Causal discovery in python. *JMLR*, 25(60):1–8, 2024.
- [29] H. Gweon and L. Schulz. Stretching to learn: Ambiguous evidence and variability in preschoolers’ exploratory play. In *CogSci*, pages 570–574, 2008.
- [30] S Kambhampati and et al. Llms can’t plan. *arXiv:2402.01817*, 2024.
- [31] C. Legare. The contributions of explanation and exploration to children’s scientific reasoning. *Child Dev. Perspect.*, 8(2):101–106, 2014.
- [32] A. Sharma and E. Kiciman. Dowhy: An end-to-end library for causal inference. *arXiv*, 2020.
- [33] C. Walker and A. Nyhout. Asking “why?” and “what if?: The influence of questions on children’s inferences. In *The Questioning Child*, pages 252–280. 2020.