IDEA: Enhancing the Rule Learning Ability of Large Language Model Agent through Induction, Deduction, and Abduction

Kaiyu He, Mian Zhang, Shuo Yan, Peilin Wu, Zhiyu Zoey Chen
Department of Computer Science
University of Texas at Dallas
{kaiyu.he, zhiyu.chen2}@utdallas.edu

Abstract

While large language models (LLMs) have been thoroughly evaluated for deductive and inductive reasoning, their proficiency in abductive reasoning and holistic rule learning in interactive environments remains less explored. We introduce RULEARN, a novel benchmark specifically designed to assess the rule-learning abilities of LLM agents in interactive settings. In RULEARN, agents strategically interact with simulated environments to gather observations, discern patterns, and solve complex problems. To enhance the rule-learning capabilities for LLM agents, we propose IDEA, a novel reasoning framework that integrates the process of **I**nduction, **DE**duction, and **A**bduction. The IDEA agent generates initial hypotheses from limited observations through abduction, devises plans to validate these hypotheses or leverages them to solve problems via deduction, and refines previous hypotheses using patterns identified from new observations through induction, dynamically establishing and applying rules that mimic human rule-learning behaviors. Our evaluation of the IDEA framework, which involves five representative LLMs, demonstrates significant improvements over the baseline. Furthermore, within this framework, our comparison with 50 human participants reveals notable discrepancies in rule-learning behaviors. LLM agents tend to generate plausible initial hypotheses but struggle to refine them through interaction. Conversely, humans, despite sometimes overlooking initial details, excel at incorporating feedback and continuously improving their hypotheses. We believe our benchmark, RULEARN, will serve as a valuable and challenging resource, and that the IDEA framework will provide crucial insights for the development of LLM agents capable of human-like rule learning in real-world scenarios.[1]

1 Introduction

One major pillar of human intelligence is the ability to discern rules and apply them. We identify patterns, formulate hypotheses, and refine them by interacting with the environment. This exploratory process traditionally involves three stages: abduction, deduction, and induction. According to Charles Peirce's definition [Frankfurt, 1958, Peirce, 1974], the rule-learning loop typically begins with an explanatory hypothesis that arises from **abduction**. This is followed by iterative experiments guided by the hypothesis, known as **deduction**, which leads to the modification and refinement of the hypothesis through **induction** (As shown in Figure 2). Whether the hypothesis emerges from rigorous logical reasoning or a more creative process is debated, but the cyclic nature of this reasoning is pivotal in advancing our understanding and problem-solving capabilities in

[1]Code and data: https://github.com/MeanStudent/RULEARN_IDEA_project

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

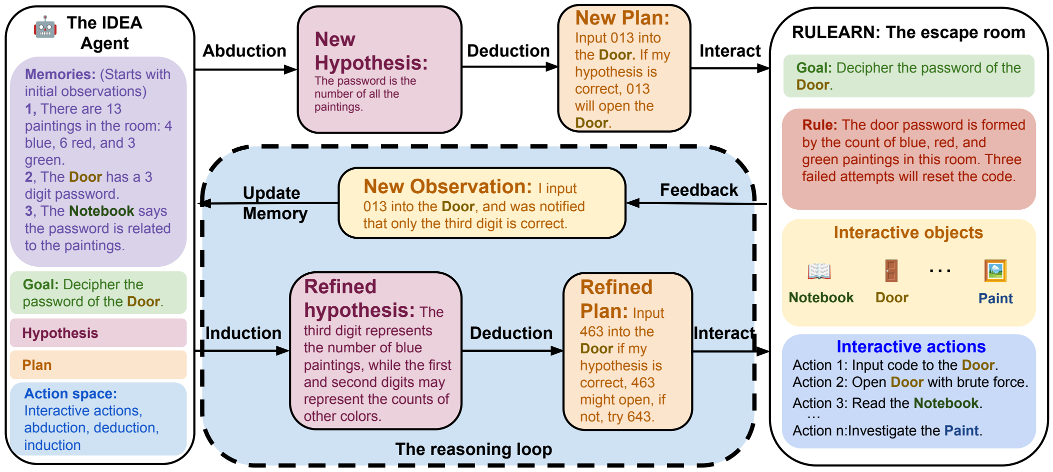

Figure 1: A simplified puzzle in the RULEARN benchmark and the IDEA agent's workflow (in real puzzles, agents have fewer initial observations and more complex rules). The agent generates an initial hypothesis through abduction, develops an exploration plan via deduction, and refines its hypothesis using induction. For example, the IDEA agent first hypothesizes that the password is the number of the blue paintings, tests this by entering code 003, and adjusts its strategy based on the results.

various disciplines. This establishes the foundation of how humans learn the rules of the world and apply them to effect change.

For such cyclic reasoning to function effectively, it requires an interactive foundation. In the real world, the observations we reason from are gathered through interaction. Then, through alternating cycles of refining hypothesis and interaction, we iteratively collect evidence and refine our hypotheses, drawing ever closer to the truth. The dynamic interplay between interaction and different stages of the reasoning loop within an interactive environment is the key to humans' reasoning effectiveness. However, recent studies focusing on reasoning often lack this dynamic and interdependent nature, as abduction, induction, and deduction are tested one at a time in non-interactive environments [Bowen et al., 2024, Wang et al., 2023, Saparov et al., 2024, Liu et al., 2024].

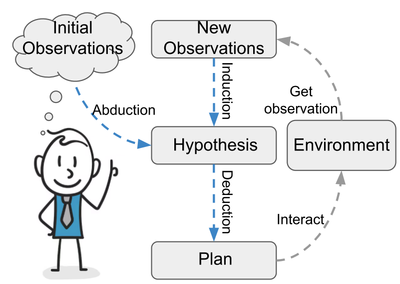

Figure 2: The reasoning cycle of rule learning encompasses abduction, deduction, and induction.

To address this gap, we introduce RULEARN, a benchmark designed to evaluate the rule-learning abilities of LLM agents in interactive environments. RULEARN consists of puzzles with hidden rules set in a text-based environment, where agents begin exploration without prior knowledge of these rules. The agent's objective is to solve the puzzle by interacting with the environment. At each step, the agent selects an action from a predefined set, executes it, and gathers observations to gain insights into the environment. Successfully solving the puzzles requires the agent to strategically select actions, efficiently gather pattern-revealing observations, and accurately reason from them to infer the hidden rules. RULEARN presents substantial challenges, as agents must rely entirely on observations generated by their chosen actions to discern rules. If these observations fail to reveal a clear pattern, the agent is likely to fail.

We design three types of environments within RULEARN to evaluate the rule-learning ability in different scenarios: (1) **The Function Operator:** The agent is tasked with determining the coefficients of mathematical functions defined by hidden expressions. Agents can assign various values to the input variables and observe the outputs, using this information to hypothesize the function's form. The challenge lies in efficiently selecting input values that reveal the underlying structure and accurately computing the coefficients based on limited observations. (2) **The Escape Room:** The agent is tasked with deciphering the passcode to exit an escape room. A hidden rule determines how the objects in the room infer the passcode digits. Agents interact with these objects to gather clues and input passcodes into the door. The door provides hints indicating which digits are correct. Based on these hints, agents formulate hypotheses and infer the relationship between the objects and the passcode. (3) **The Reactor:** The agent is tasked with synthesizing target strings using a reactor with a hidden string-combining rule. Agents need to experiment with different inputs and analyze outputs to deduce the reactor's transformation rule and achieve the desired outcome. Detailed puzzle definitions are provided in § 3.

To tackle the challenge in RULEARN, we introduce IDEA, a novel reasoning framework that integrates the process of **I**nduction, **DE**duction, and **A**bduction. The IDEA agent employs these reasoning processes iteratively to explore the environments, learn rules, and achieve goals. In the **abduction** phase, the IDEA agent generates an initial hypothesis from limited observations. During the **deduction** phase, the IDEA agent creates and executes plans to meet objectives or test its hypothesis. In the **induction** phase, the IDEA agent refines its hypothesis based on new observations, enhancing their accuracy and robustness. This iterative cycle enables the LLM agent to continually improve the learned rules through environmental feedback. An overview of how the IDEA agent solves puzzles in RULEARN is shown in Figure 1.

We evaluat the IDEA framework using five popular LLMs—GPT-3.5-Turbo, GPT-4o, Gemma-7B, Llama3-8B and Llama3-70B. The experimental results demonstrate that IDEA significantly improves rule-learning proficiency and puzzle-solving performance of the LLM agent within the RULEARN benchmark. The IDEA agent increases success rates by approximately 10% compared to the baseline LLM agent. Specifically, the action of the IDEA agent aims to validate or leverage current hypothesis, providing important guidance in an environment with hidden rules. Without the guidance of the hypothesis, the LLM agent will always choose the most direct actions attempting to quickly solve the puzzle. Such actions, taken without understanding the hidden rules, are akin to navigating without a compass—efforts are made, but they often lead to ineffective or misguided outcomes. Based on our experiments, the IDEA agent achieves 47.6% fewer repeated actions, yielding more diverse observations and finally solving the puzzles with a better understanding of the hidden rule. These improvements suggest that IDEA enables LLMs to better adapt and learn rules in unfamiliar environments, narrowing the gap between human and AI rule-learning abilities.

Within the IDEA framework, we conduct a study to compare the rule-learning behaviors and performances between LLMs and 50 human participants solving puzzles. Our findings reveal that, despite the improvements brought by the IDEA framework, LLMs still face the following challenges when compared to human participants: (1) LLM agents continue to struggle with efficiently exploring unfamiliar environments. The new observations gathered usually do not provide enough information to help reveal the underlying rules, moreover, LLM agents tend to repeat previous actions that do not yield new information. (2) LLMs fall short in deducting valid plans to verify current hypotheses and guide future exploration. (3) LLMs fall short to correct previous hypotheses when they contradict new observations and are less capable of refining a hypothesis to make it more robust and encompassing all observations. These findings provide valuable insights for the development of LLM agents capable of human-like rule learning in real-world scenarios.

2 Related Works

2.1 LLM agent

Agents driven by large language models equipped with memory, tool use, planning, and reasoning capabilities have shown significant advancements in understanding complex tasks and achieving notable performance [Wang et al., 2024b]. Techniques like memory summarization [Chen et al., 2023, Zhou et al., 2023, Wang et al., 2024a] and retrieval [Andreas, 2022, Park et al., 2023, Zhong et al., 2023] assist LLM agents in filtering relevant information based on the current context [Zhang et al., 2024]. Additionally, the integration of tool use enables LLM agents to perform better in knowledge-grounded tasks. For instance, some studies [Nakano et al., 2022, Lu et al., 2023, Shi et al., 2023a] have enabled LLM agents to access up-to-date information from external databases to provide correct answers, while others [Schick et al., 2023] have facilitated the use of various tools like calculators to enhance solution accuracy. Planning modules [Yuan et al., 2023, Shen et al., 2023] allow these agents to decompose complex tasks into simpler components, achieving significant improvements. Moreover, employing search algorithms [Yao et al., 2023, Besta et al., 2024] helps LLM agents reach solutions in fewer steps. Building on this foundation, we are interested in whether LLM agents can derive new knowledge–learning rules through interacting with the environment, rather than merely answering questions. Recent studies have begun investigating different types of reasoning in LLMs, including abduction, deduction, and induction [Bowen et al., 2024, Wang et al., 2023, Saparov et al., 2024, Cheng et al., 2024, Yang et al., 2024]. However, existing research mostly focuses on isolated aspects of the rule-learning process. Consequently, the comprehensive rule-learning abilities of LLMs in interactive environments remain largely underexplored, a topic that will be further examined in Section 2.2.

2.2 Rule Learning

Rule learning—involving evidence gathering, pattern identification, hypothesis formulation, and validation—is gaining focus in recent research. Yang et al. [2023] and Zhu et al. [2024] improved question-answering systems by formulating optimal response rules, while Shi et al. [2023b] advanced event prediction by enabling LLMs to hypothesize causes of previous events. Some works also evaluate LLMs rule-learning abilities [Liu et al., 2024]. However, existing studies adopt static, non-interactive settings for rule learning, where LLMs are provided with complete information to formulate rules. This approach has two significant limitations: it contrasts with real-world rule learning, which involves actively and strategically gathering evidence, and it restricts agents from conducting experiments to validate the rules they learn. In real-world scenarios, the process of gathering observations, validating hypotheses, and refining hypotheses is iterative. These simplified, static settings fail to adequately assess an agent's capabilities in real, dynamic environments. Even in studies with interactive environments [Xu et al., 2024, Montes et al., 2022], information gathering, rule generation, and application are treated as independent objectives or are incomplete. In reality, these processes are dynamic and interdependent: generated hypotheses guide actions, and new observations refine those hypotheses. In contrast, RULEARN provides a realistic setting that allows LLMs to engage in dynamic, interdependent rule-learning processes, closely mirroring real-world scenarios. To effectively navigate and learn within this environment, IDEA equips LLM agents with the ability to emulate human reasoning by managing these interdependent processes.

3 The RULEARN Benchmark

We develop three puzzle sets, the **Function Operator**, the **Escape Room**, and the **Reactor**, each comprising 20 puzzles associated with environments of varying complexity.

The Function Operator. This puzzle type simulates scenarios where systemic theories or established knowledge (in this case, mathematics) is applicable. The agent interacts with a set of univariate multi-term equations involving integer parameters from [0,9] and elementary functions of the variable x, selected from $f(x) \in \{x^0, x^1, x^2, \sin(x), \frac{1}{x}, |x|, -x\}$. The agent is provided with the number of functions, the presence of specific parameters in each function (the exact numerical values of these parameters are unknown and represented by letters), and the types of elementary functions involved in the current puzzle. The goal of the agent is to deduce the values of these parameters. For example, in one puzzle, the ground truth is $\mathbf{F_1}(x) = a\sin(x) + b \times \frac{1}{x}$, $\mathbf{F_2}(x) = ax^2$ where $a = 3$ and $b = 2$. The agent knows the following information: There are three elementary functions in this puzzle $\{\sin(x), \frac{1}{x}, x^2\}$, there are two functions $\mathbf{F_1}(x)$ and $\mathbf{F_2}(x)$, $\mathbf{F_1}(x)$ has 2 terms and parameters a, b in it, and $\mathbf{F_2}(x)$ has 1 term and one parameter a. To solve the puzzle, the agent must interact with the environment through a defined action space: selecting a function and assigning values to x, then observing the resultant output. For example, assigning values 1 and 2 to $\mathbf{F_2}$ reveals a quadratic increase in output, indicating the presence of x^2 in $\mathbf{F_2}$. Similarly, assigning a value of 1 to $\mathbf{F_1}$ results in a floating-point output, rather than an integer, suggesting the inclusion of trigonometric components, confirming that $\sin(x)$ is a component of $\mathbf{F_1}$. Each puzzle features a different number of functions, unknown parameters, and elementary functions in use.

The Escape Room. This environment simulates scenarios where no established knowledge is applicable. We create a fictitious environment: an agent is in an escape room–an art gallery, tasked with deciphering a password to unlock a code-secured door and escape. The password is a 3-digit number where each digit represents the count of paintings of a specific **type**—watercolor, oil, or acrylic—that share a given **color**. The agent can explore the environment and investigate paintings, receiving brief descriptions such as *This is an acrylic painting of a green jungle*, providing information about the painting's type and color. The agent only knows that the password is a 3-digit number

Table 1: The reacting rules in the Reactor Puzzle. All letters are functionally equivalent and exhibit no special behaviors. Identical symbols represent the same letter, while different symbols denote different letters. Each puzzle operates under one specific rule. The Middle Insertion rule inserts the shorter string into the longer string; if the length of the longer string is odd, the shorter string is inserted just to the right of the center. If both strings are of equal length, the second string is inserted into the middle of the first string. The Prefix Replacement rule retains the prefix of the longer string and concatenates it with the shorter string, dropping the tail of the longer string results in two output strings. There are two special cases where the strings are simply concatenated in order.

Rule Description	Example Reaction 1	Example Reaction 2	Special Case 1	Special Case 2
Simple Concatenation	AB + C = ABC	AB + CDE = ABCDE	—	—
Reverse Concatenation	AB + C = CAB	AB + CDE = CDEAB	—	—
Middle Insertion	AB + C = ACB	AB + CDE = CDABE	A + B = AB	—
Prefix Replacement	AB + C = AC + B	AB + CDE = CAB + DE	AB + CD = ABCD	AA + A = AAA

and receives a hint to focus on a specific color relevant to the password. After forming an initial hypothesis and inputting the password into the door, the door provides hints indicating which digits are correct. Based on these hints, the agent refines its hypotheses and infers the relationship between the painting attributes and the passcode. To prevent the problem from being solved merely by brute force, the specific color associated with the password changes after every three failed attempts. Each puzzle presents a varying number of paintings; paintings in the same room as the agent are directly visible, while those in different rooms are not, requiring the agent to take actions such as moving to another gallery to reveal all the paintings.

The Reactor. This environment type also simulates scenarios where no established knowledge is applicable. However, unlike in the Escape Room where agents only input passwords to test hypotheses, in the Reactor, agents are challenged to design more fine-grained experiments to uncover hidden rules. Specifically, the agent's task is to synthesize target strings using a reactor governed by a hidden string-combining rule. These strings are represented by sequences of alphabetic letters, such as *A*, *B*, *AABB*, and *CAB*. The reactor permits the agent to input two strings, initiating a reaction that produces a new string for use in subsequent experiments. The agent's objective is to decipher the specific rules that govern string synthesis by methodically testing different string combinations, with the ultimate goal of synthesizing the target string using the discovered rules. We have designed four types of rules, detailed in Table 1.

The detailed statistics for each puzzle type and example puzzles are provided in Appendix B.2 and B.5, respectively. Note that the RULEARN benchmark is designed to simulate real-world interactions and is adaptable to complex text environments with diverse rules. To achieve this, we do not restrict the representation of rules to a specific formal language, as doing so would limit realism and complexity. Moreover, predefining the rule format would provide agents with prior knowledge, contradicting real-world conditions. Therefore, we enable the LLM agents to use natural language to describe rules in RULEARN, promoting generalizability and future extensions of the benchmark.

4 The IDEA Agent

We introduce IDEA, a novel reasoning framework that integrates the process of **I**nduction, **DE**duction, and **A**bduction to learn rules in interactive environments. The IDEA agent consists of the following components: Goal(G), Action Space($\mathbb{A}$), Memory($\mathbb{M}$), Hypothesis(H), and Plan(P), which are elaborated in Table 2.

Upon beginning to explore a puzzle, we initialize the agent memory with an initial observation of the environment. The agent's goal is initialized with the objective of the puzzle, e.g., *synthesize a target string* for a Reactor puzzle. The agent's action space is initialized as the set of interactive actions defined by the puzzle, such as *choosing two strings and running the Reactor*, as well as establish the initial hypothesis (abductive action), devises a plan to validate or leverage hypothesis (deductive action), and refining the current hypothesis (inductive action).

The IDEA agent begins with an abductive action to generate an initial hypothesis, followed by a deduction step to create a new plan. Based on this plan, the agent interacts with the environment. Upon receiving feedback from the environment as a new observation, the agent may take an inductive action to refine the hypothesis or perform another interaction with the environment. Deductive

Algorithm 1 IDEA Agent Rule-learning Loop

```
procedure RULELEARNINGLOOP
    Initialize Goal(G), Action Space(𝔸)
    Memory(𝕄) ← Initial observations
    #step ← 0
    Hypothesis(H) ← Abduct(G, 𝔸, 𝕄)
    Plan(P) ← Deduct(H, G, 𝕄, 𝔸)
    𝕄.add("New hypothesis and plan", H, P)
    while G not achieved and #step ≤ max_step do
        a ← select_action(G, H, P, 𝕄, 𝔸)
        if a is interactive action then
            result ← execute_action(a, G, H, P, 𝕄)
            𝕄.add(result)
            #step ← #step + 1
        else if a is inductive action then
            H ← Induct(a, G, 𝕄, H, P)
            P ← Deduct(H, G, 𝕄, 𝔸)
            𝕄.add("Refined hypothesis and plan", H, P)
        end if
    end while
end procedure
```

The IDEA Agent Component	Definition
Goal(G)	Goal of the agent in the current puzzle.
Action Space($\mathbb{A}$)	Set of actions the agent can take, including abductive action, deductive action, inductive action, as well as the set of interactive actions defined by the puzzle.
Memory($\mathbb{M}$)	Set of natural language strings to record all interaction results till the current step.
Plan(P)	Generated plans to guide future actions.

Table 2: Components of the IDEA agent.

action is invoked to adjust the plan every time the hypothesis changes. This reasoning loop continues until the puzzle is solved or a maximum number of steps is reached. After each step, the results are appended to the agent's memory, including interaction outcomes and any modifications to the hypothesis or plan. We provide a simplified algorithm demonstrating how the IDEA agent operates in Algorithm 1. Specifically, at each step, we prompt the LLM to reflect on the information recorded in the IDEA agent's components to make decisions and take actions. We employ the chain-of-thought (COT) reasoning [Wei et al., 2022] for all prompts. Detailed prompts for each type of action are available in Appendix B.4. More detailed implementation of the agent can be found in Appendix B.3.1. Similar to real-life scenarios, when agents solve tasks in RULEARN puzzles, they do not know the outcomes in advance. Consequently, it is challenging to decide when to refine or change their hypothesis and plans, as well as what interactive actions to take to gather pattern-revealing observations. A detailed example of the IDEA agent solving the Reactor Puzzle is provided in Figure 3.

4.1 Experiment Settings

To evaluate the effectiveness of IDEA, we respectively initialize it with three popular open-source LLMs, including Gemma-7B [Team et al., 2024], Llama3-8B, and Llama3-70B [Dubey et al., 2024], and two closed-source LLMs, GPT-3.5-Turbo [Dubey et al., 2024] and GPT-4o [2]. We compare the IDEA agent against the following two variants:

- **Baseline Agent**: The Baseline agent lacks the reasoning loop of abduction, deduction, and induction and does not generate hypotheses or plans. At each step, it selects an interactive action solely based on its current memories and the goal.
- **Oracle-rule Agent**: Even if the agent could successfully learn the correct rule, applying the learned rule to solve the puzzle is non-trivial. The Oracle-rule agent serves as a special control group to establish the Oracle performance with the ground-truth rule provided at the beginning. Specifically: 1) For the Function Operator puzzles, agents are given the exact forms of the functions. Their task is to derive the values of the coefficients. 2) For the Escape Room puzzles, agents are provided with how the password is constructed from the objects in the room. Their task is to derive the password using the provided rule. 3) For the Reactor puzzles, the reaction rule is given to the agents in natural language accompanied with examples. The agents only need to synthesize the target strings using the provided rules.

[2] https://openai.com/index/gpt-4o-system-card/

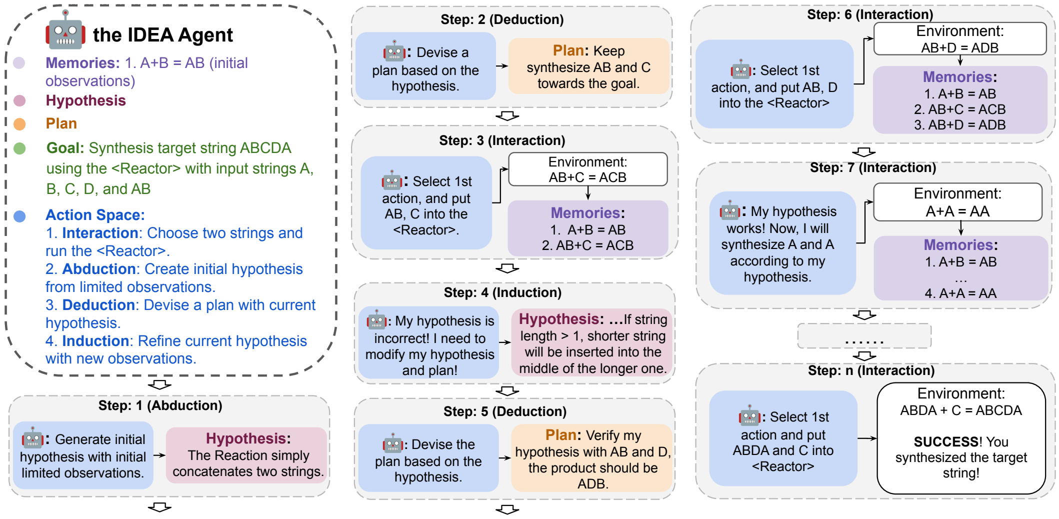

Figure 3: An example of the IDEA agent solving a Reactor puzzle. At each step, the agent must choose whether to interact with the environment or adjust its hypothesis and plan based on current observations. If observed facts contradict the existing hypothesis, the agent is expected to refine its hypothesis. The refined hypothesis and plan will then guide subsequent exploration.

Each variant is evaluated on all three puzzle types (totaling 60 puzzles), with five trials per puzzle, resulting in 300 tests per variant. In our initial experiments, we found that using a higher temperature enhanced the agent's creativity, allowing it to generate more diverse and higher-quality hypotheses that aided in discovering the underlying rules. Therefore, we set the temperature to 1.0. We apply the COT for the prompting of each step for all agents.

Additionally, since we observed that the success rate does not improve after 15 interactive steps, we set the maximum interaction step count to 15. The agent is considered to have failed a puzzle if it doesn't solve it within 15 interactive steps.

4.2 Human participants

To compare the behavior and performance of humans with that of LLMs in the abduction, deduction, and induction reasoning processes, we recruited 50 participants and randomly selected 30 puzzles (10 from each of the three puzzle types), assigning each participant three puzzles—one from each type. Each puzzle was attempted by five different individuals, ensuring no prior knowledge of the underlying rules and maintaining an unbiased sample across puzzle types. Details about Institutional Review Board (IRB) approval and participant recruitment are provided in §A. The human participants are instructed to operate under the same reasoning procedure as IDEA.

Failure to solve a puzzle within 15 steps resulted in marking the attempt as unsuccessful, ensuring a fair comparison.

4.3 Main Results

We calculated the Wilson Confidence Interval (CI) [Brown et al., 2001] for puzzle solving success rate across different variants. The detailed results are displayed in Table 3.

For the Oracle-rule agent, GPT-4o exhibits superior deductive capabilities, outperforming other models with an approximate 50% success rate across all three puzzle types. Specifically, in the Escape Room puzzles, agents achieve up to an 80% success rate by simply following the provided rule to count the number of paintings of specific colors and types. However, in the Function Operator and Reactor puzzles, merely knowing the rule is not sufficient for success, applying the rule to solve the puzzle is more challenging. Agents must also devise effective plans to apply these rules to reach a solution. This additional planning requirement results in a comparatively lower success rate in these puzzles.

For the Baseline agent, where they are not provided with the underlying rules and solely rely on historical observations to make interactive actions, GPT-4o continues to lead in performance. Notably,

Table 3: Puzzle Success Rate (95% CI). The table displays the success rate for each setting, with each pair of numbers in parentheses representing the lower and upper bounds of the 95% confidence interval for the success rate. The human results are based on a randomly selected 50% subsample of the puzzles. We can see that across all LLMs, IDEA achieves consistent significant improvements except for the Reactor puzzles. We use **Boldface** to highlight the performance comparisons between Baseline and IDEA Agents with GPT-4o.

Setup	LLMs	All Types (%)	Function Operator (%)	Escape Room (%)	Reactor (%)
Oracle-rule Agent	Gemma-7B	(0.71, 3.84)	(0.0, 3.7)	(1.57, 9.84)	(0.18, 5.45)
	Llama3-8B	(7.09, 13.92)	(2.78, 12.48)	(12.51, 27.78)	(2.15, 11.18)
	Llama3-70B	(33.33, 44.29)	(28.18, 46.78)	(48.21, 67.2)	(14.17, 29.98)
	GPT-3.5-Turbo	(3.83, 9.28)	(0.55, 7.0)	(4.11, 15.0)	(4.11, 15.0)
	GPT-4o	(47.35, 58.57)	(35.61, 54.76)	(72.22, 87.49)	(24.56, 42.69)
Baseline Agent	Gemma-7B	(0.34, 2.9)	(0.0, 3.7)	(0.18, 5.45)	(0.55, 7.0)
	Llama3-8B	(2.55, 7.27)	(0.55, 7.0)	(4.81, 16.23)	(0.55, 7.0)
	Llama3-70B	(18.6, 28.09)	(22.78, 40.63)	(12.51, 27.78)	(12.51, 27.78)
	GPT-3.5-Turbo	(3.57, 8.89)	(5.52, 17.44)	(1.03, 8.45)	(1.57, 9.84)
	GPT-4o	**(19.21, 28.79)**	**(18.6, 28.09)**	**(14.17, 29.98)**	**(14.17, 29.98)**
IDEA Agent (Ours)	Gemma-7B	(0.06, 1.86)	(0.18, 5.45)	(0.0, 3.7)	(0.0, 3.7)
	Llama3-8B	(1.36, 5.17)	(0.18, 5.45)	(3.43, 13.75)	(0.0, 3.7)
	Llama3-70B	(25.41, 35.76)	(34.67, 53.77)	(21.89, 39.58)	(10.89, 25.55)
	GPT-3.5-Turbo	(10.53, 18.38)	(22.78, 40.63)	(2.78, 12.48)	(2.15, 11.18)
	GPT-4o	**(30.14, 40.9)**	**(30.94, 49.8)**	**(32.8, 51.79)**	**(16.69, 33.23)**
	Human	(55.38, 70.62)	(52.15, 77.56)	(42.31, 68.84)	(54.19, 79.24)

Llama3-70B narrows the performance gap, especially in the Function Operator puzzles. Across models, the success rates drop by about half compared to the Oracle-rule agent. This significant decrease highlights the challenge of rule learning and indicates that current LLM agents struggle to learn rules in unfamiliar environments without explicit guidance.

The IDEA Agent significantly boosts success rates. Our proposed IDEA framework leads to approximate 10% increases in success rates for Llama3-70B, GPT-3.5-Turbo, and GPT-4o compared to the Baseline agent. This improvement demonstrates that incorporating a reasoning loop of abduction, deduction, and induction substantially enhances the LLM rule-learning performance in unfamiliar environments. IDEA enables the LLMs to generate hypotheses, plan actions, and refine their understanding based on new observations, which is crucial for rule learning. However, smaller models like Llama3-8B and Gemma-7B experience a slight decline in performance when applying IDEA. This decline is likely due to the increased complexity and context management demands introduced by the IDEA framework, which pose challenges for small models with limited capacity.

LLM agents still fall far behind humans. In the Escape Room puzzle, where the primary challenge is to discover the rule, the Oracle-rule agents excel because once the rule is identified, applying the rule is simple. However, in other types of puzzles, human participants significantly outperform all LLM agents, including the Oracle-rule agents, even without knowing the rules beforehand. This notable disparity highlights that LLM agents lag substantially behind humans, not only in learning new rules but also in effectively applying known rules to solve problems.

4.4 Analysis

IDEA boosts puzzle-solving speed. Figure 4 illustrates the cumulative number of puzzles solved at each interaction step for the Baseline agent, the IDEA agent, and human participants. The slopes of the lines represent the rate at which puzzles are solved per step. Compared to the Baseline agent, the IDEA agent exhibits a steeper slope, indicating that the integration of abductive, deductive, and inductive reasoning enhances the agent's efficiency in exploring the environment and learning the underlying rules, especially during the early stages.

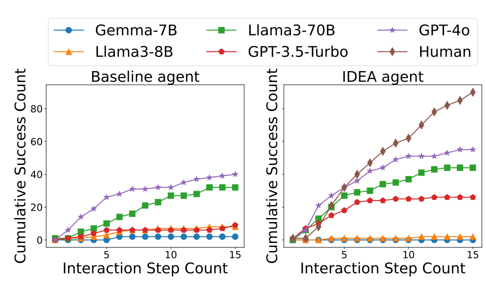

Figure 4: Comparison of the cumulative number of puzzles solved at each interaction step. The IDEA agent significantly increases the solving speed compared to the Baseline agent.

When focusing on human participants, we observe that they solve fewer puzzles in the initial steps compared to the LLM agents. However, as interactions continue, the number of puzzles solved by humans increases rapidly, eventually surpassing that of all LLM agents. In contrast, LLM agents solved 88.76% of the puzzles within the first 10 steps. Beyond this point, additional interactions contribute less to their success rate. This pattern suggests that humans have a superior ability to learn continuously from interactive environments, effectively improving their performance over time. If we did not limit the puzzles to 15 steps, we anticipate that the success rate of human participants would be even higher. We will explore this observation further in Section 5.

IDEA reduces repetitive actions. LLM agents frequently repeat previous actions instead of exploring new ones. This behavior is highly inefficient in our controlled puzzle environments, where each interaction yields deterministic results, and repeating the same action generally does not provide new information. An exception is the Escape Room puzzle, where entering the same password could result in different outcomes after the rule changes following three failed attempts.

We calculate the average number of repeated actions performed while solving each puzzle, with detailed statistics presented in Table 7 in Appendix B.2. We observe that most LLMs commonly repeat actions in the Baseline agent. Although models like GPT-4o and Llama3-70B exhibit this issue to a lesser degree, even these advanced models sometimes repeat actions that fail to provide further insights. The IDEA agent effectively reduces this tendency by explicitly generating plans during the deduction phase. By outlining a clear plan, the IDEA agent can better assess whether the current observations are sufficient or if further specific evidence is needed to reveal the underlying rule. For example, in the Escape Room puzzle, the IDEA agent avoids unnecessary attempts at entering passwords when the evidence gathered is sufficient to determine the correct code (see Figure 11 in Appendix B.4).

IDEA relies on the reasoning ability of underlying LLMs. The effectiveness of IDEA depends on the underlying LLMs' ability to reason from hypotheses and observations. Particularly, if an agent generates a false hypothesis and fails to properly refine it, being guided by this incorrect hypothesis can lead the agent to perform even worse than the baseline. During our experiments, we observed that current LLMs tend to hallucinate, especially in the Escape Room puzzles and more severely in the Reactor puzzles. This results in smaller performance improvements compared to those seen with the Function Operator puzzles. This is likely because such fictitious scenarios are not extensively represented in LLM training data. Moreover, LLMs struggle to recognize letter-level patterns, and their reasoning capabilities still require significant enhancement. Examples of hallucination can be seen in Appendix B.4.4).

5 Human evaluation

We employ three computer science graduate students (our co-authors) to evaluate all the hypotheses and plans generated by both the IDEA agent and human participants during the abduction, deduction, and induction stages within a randomly selected 50% subsample of the puzzles.

Abduction stage. During the abduction stage, agents formulate an initial hypothesis based on initial observations. Since some puzzles are simple, agents can potentially guess the ground truth rule at this stage. Figure 5(a) shows that LLMs such as GPT-4o have approximately a 30% success rate in correctly identifying the rule during abduction. Notably, humans exhibit the lowest success rate in this stage. We believe this is because humans often lack the patience to thoroughly review all textual information initially and may overlook certain details. In contrast, LLMs meticulously process every word of the prompt and leverage their comprehensive pre-training data to generate plausible hypotheses. Additionally, humans tend not to formalize hypotheses when operating under uncertainty.

Deduction stage. After establishing an initial hypothesis—or each time the agent refines a hypothesis—the agent derives a plan to either validate this hypothesis or use it to solve the puzzle. As shown in Figure 5(b), humans generally outperform LLMs in creating high-quality plans. However, in the Function Operator puzzles, Llama3-70B and GPT-4o perform better than humans, suggesting that these models may have superior mathematical intuition compared to our human participants, who were undergraduate students from computer science-related majors.

Induction stage. Figure 5(c) illustrates the proportion of inductions where the refined hypothesis improved over the previous one, which we define as the effective induction rate. Humans outper-

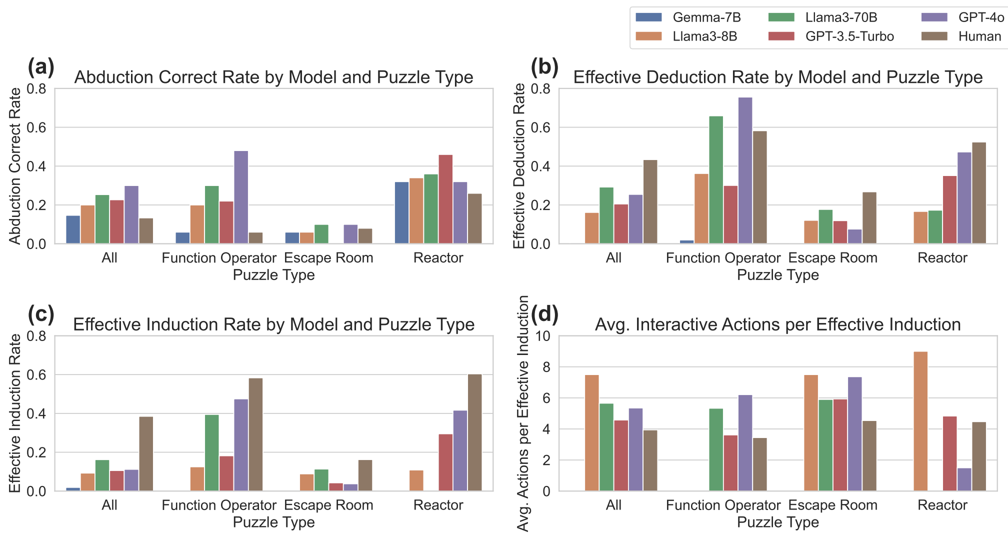

Figure 5: Human Evaluation Results. Bars represent measured values per model and puzzle type; the absence of a bar indicates zero or unavailable data. Plot (a): Abduction Correct Rate—the frequency of correctly guessing the rule during abduction. Plot (b): Effective Deduction Rate—the rate at which deduction plans effectively validate hypotheses or solve puzzles. Plot (c): Effective Induction Rate—the proportion of inductions where the refined hypothesis improved over the previous one. Plot (d): Average Actions per Effective Induction—the average number of interactive actions needed for an effective induction.

form LLMs by a significant margin in this process, with 40% of their refined hypotheses being better than the previous ones. Surprisingly, GPT-4o performs worse than Llama3-70B and GPT-3.5-Turbo. We found that GPT-4o tends to hallucinate, especially in Escape Room puzzles, which negatively impacts its overall effective induction rate. In contrast, Llama3-70B almost never engages in induction activities within Reactor puzzles, as it fails to recognize when its hypothesized rules contradict its observations (see Appendix B.4.4). Despite these issues, both Llama3-70B and GPT-4o exhibit performance closer to humans during the induction process compared to other models in Function Operator puzzles. This relative success is believed to be due to the extensive training of LLMs with mathematical content, which ensures stable and superior performance in these types of puzzles. In contrast, Escape Room and Reactor puzzles, which are less likely to appear extensively in training datasets, see poorer and more variable performance from the models.

Average interactions needed for effective induction. Figure 5(d) shows that humans require fewer interactions—approximately four on average—to effectively refine their hypotheses, compared to LLMs. While LLMs can process initial information thoroughly and generate plausible hypotheses, they face challenges in refining these hypotheses based on new observations during interaction with the environment (see Figure 7 in Appendix **??**). This limitation suggests that LLMs may struggle to learn from new observations and incorporate feedback to continuously improve their hypotheses and problem-solving strategies. This gap may become more pronounced when agents are faced with larger action spaces and environments governed by more complex rules.

6 Conclusion

In this work, we introduce RULEARN, a benchmark designed to evaluate LLM agents' rule-learning abilities in interactive environments. We propose IDEA, an agent framework that mimics human reasoning through abduction, deduction, and induction. Comprehensive experiments involving five prominent LLMs and human participants reveal that while IDEA significantly improves the rule learning ability of LLM agents, there is still a large gap between LLM and humans particularly in refining hypotheses and adapting strategies. Despite these advancements from the IDEA framework, LLMs still face challenges in generating valid hypotheses and avoiding repetitive actions in complex scenarios. Our findings underscore the need for further development of LLMs that can emulate human cognitive processes more effectively in explorations of novel environments. RULEARN provides a foundational resource for future research aimed at closing these gaps.

References

Jacob Andreas. Language models as agent models, 2022. URL https://arxiv.org/abs/2212.01681.

Maciej Besta, Nils Blach, Ales Kubicek, Robert Gerstenberger, Michal Podstawski, Lukas Gianinazzi, Joanna Gajda, Tomasz Lehmann, Hubert Niewiadomski, Piotr Nyczyk, and Torsten Hoefler. Graph of thoughts: Solving elaborate problems with large language models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(16):17682–17690, March 2024. ISSN 2159-5399. doi: 10.1609/aaai.v38i16.29720. URL http://dx.doi.org/10.1609/aaai.v38i16.29720.

Chen Bowen, Rune Sætre, and Yusuke Miyao. A comprehensive evaluation of inductive reasoning capabilities and problem solving in large language models. In *Findings of the Association for Computational Linguistics: EACL 2024*, pages 323–339, 2024.

Lawrence D Brown, T Tony Cai, and Anirban DasGupta. Interval estimation for a binomial proportion. *Statistical science*, 16(2):101–133, 2001.

Howard Chen, Ramakanth Pasunuru, Jason Weston, and Asli Celikyilmaz. Walking down the memory maze: Beyond context limit through interactive reading, 2023. URL https://arxiv.org/abs/2310.05029.

Kewei Cheng, Jingfeng Yang, Haoming Jiang, Zhengyang Wang, Binxuan Huang, Ruirui Li, Shiyang Li, Zheng Li, Yifan Gao, Xian Li, Bing Yin, and Yizhou Sun. Inductive or deductive? rethinking the fundamental reasoning abilities of llms, 2024. URL https://arxiv.org/abs/2408.00114.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, Anirudh Goyal, Anthony Hartshorn, Aobo Yang, Archi Mitra, Archie Sravankumar, Artem Korenev, Arthur Hinsvark, Arun Rao, Aston Zhang, Aurelien Rodriguez, Austen Gregerson, Ava Spataru, Baptiste Roziere, Bethany Biron, Binh Tang, Bobbie Chern, Charlotte Caucheteux, Chaya Nayak, Chloe Bi, Chris Marra, Chris McConnell, Christian Keller, Christophe Touret, Chunyang Wu, Corinne Wong, Cristian Canton Ferrer, Cyrus Nikolaidis, Damien Allonsius, Daniel Song, Danielle Pintz, Danny Livshits, David Esiobu, Dhruv Choudhary, Dhruv Mahajan, Diego Garcia-Olano, Diego Perino, Dieuwke Hupkes, Egor Lakomkin, Ehab AlBadawy, Elina Lobanova, Emily Dinan, Eric Michael Smith, Filip Radenovic, Frank Zhang, Gabriel Synnaeve, Gabrielle Lee, Georgia Lewis Anderson, Graeme Nail, Gregoire Mialon, Guan Pang, Guillem Cucurell, Hailey Nguyen, Hannah Korevaar, Hu Xu, Hugo Touvron, Iliyan Zarov, Imanol Arrieta Ibarra, Isabel Kloumann, Ishan Misra, Ivan Evtimov, Jade Copet, Jaewon Lee, Jan Geffert, Jana Vranes, Jason Park, Jay Mahadeokar, Jeet Shah, Jelmer van der Linde, Jennifer Billock, Jenny Hong, Jenya Lee, Jeremy Fu, Jianfeng Chi, Jianyu Huang, Jiawen Liu, Jie Wang, Jiecao Yu, Joanna Bitton, Joe Spisak, Jongsoo Park, Joseph Rocca, Joshua Johnstun, Joshua Saxe, Junteng Jia, Kalyan Vasuden Alwala, Kartikeya Upasani, Kate Plawiak, Ke Li, Kenneth Heafield, Kevin Stone, Khalid El-Arini, Krithika Iyer, Kshitiz Malik, Kuenley Chiu, Kunal Bhalla, Lauren Rantala-Yeary, Laurens van der Maaten, Lawrence Chen, Liang Tan, Liz Jenkins, Louis Martin, Lovish Madaan, Lubo Malo, Lukas Blecher, Lukas Landzaat, Luke de Oliveira, Madeline Muzzi, Mahesh Pasupuleti, Mannat Singh, Manohar Paluri, Marcin Kardas, Mathew Oldham, Mathieu Rita, Maya Pavlova, Melanie Kambadur, Mike Lewis, Min Si, Mitesh Kumar Singh, Mona Hassan, Naman Goyal, Narjes Torabi, Nikolay Bashlykov, Nikolay Bogoychev, Niladri Chatterji, Olivier Duchenne, Onur Çelebi, Patrick Alrassy, Pengchuan Zhang, Pengwei Li, Petar Vasic, Peter Weng, Prajjwal Bhargava, Pratik Dubal, Praveen Krishnan, Punit Singh Koura, Puxin Xu, Qing He, Qingxiao Dong, Ragavan Srinivasan, Raj Ganapathy, Ramon Calderer, Ricardo Silveira Cabral, Robert Stojnic, Roberta Raileanu, Rohit Girdhar, Rohit Patel, Romain Sauvestre, Ronnie Polidoro, Roshan Sumbaly, Ross Taylor, Ruan Silva, Rui Hou, Rui Wang, Saghar Hosseini, Sahana Chennabasappa, Sanjay Singh, Sean Bell, Seohyun Sonia Kim, Sergey Edunov, Shaoliang Nie, Sharan Narang, Sharath Raparthy, Sheng Shen, Shengye Wan, Shruti Bhosale, Shun Zhang, Simon Vandenhende, Soumya Batra, Spencer Whitman, Sten Sootla, Stephane Collot, Suchin Gururangan, Sydney Borodinsky, Tamar Herman, Tara Fowler, Tarek Sheasha, Thomas Georgiou, Thomas Scialom, Tobias Speckbacher, Todor Mihaylov, Tong Xiao, Ujjwal Karn, Vedanuj Goswami, Vibhor Gupta, Vignesh Ramanathan, Viktor Kerkez, Vincent Gonguet, Virginie Do, Vish Vogeti, Vladan Petrovic, Weiwei Chu, Wenhan Xiong, Wenyin Fu, Whitney Meers, Xavier Martinet, Xiaodong Wang, Xiaoqing Ellen Tan, Xinfeng Xie, Xuchao Jia, Xuewei Wang, Yaelle Goldschlag, Yashesh Gaur, Yasmine Babaei, Yi Wen, Yiwen Song, Yuchen Zhang, Yue Li, Yuning Mao, Zacharie Delpierre Coudert, Zheng Yan, Zhengxing Chen, Zoe Papakipos, Aaditya Singh, Aaron Grattafiori, Abha Jain, Adam Kelsey, Adam Shajnfeld, Adithya Gangidi, Adolfo Victoria, Ahuva Goldstand, Ajay Menon, Ajay Sharma, Alex Boesenberg, Alex Vaughan, Alexei Baevski, Allie Feinstein, Amanda Kallet, Amit Sangani, Anam Yunus, Andrei Lupu, Andres Alvarado, Andrew Caples, Andrew Gu, Andrew Ho, Andrew Poulton, Andrew Ryan, Ankit Ramchandani, Annie Franco, Aparajita Saraf, Arkabandhu Chowdhury, Ashley Gabriel, Ashwin Bharambe, Assaf Eisenman, Azadeh Yazdan, Beau James, Ben Maurer, Benjamin Leonhardi, Bernie Huang, Beth Loyd, Beto De Paola, Bhargavi Paranjape, Bing Liu, Bo Wu, Boyu Ni, Braden Hancock, Bram Wasti, Brandon Spence, Brani Stojkovic, Brian Gamido, Britt Montalvo, Carl Parker, Carly Burton, Catalina Mejia, Changhan Wang, Changkyu Kim, Chao Zhou, Chester Hu, Ching-Hsiang Chu, Chris Cai, Chris Tindal, Christoph Feichtenhofer, Damon Civin, Dana Beaty, Daniel Kreymer, Daniel Li, Danny Wyatt, David Adkins, David Xu, Davide Testuggine, Delia David, Devi Parikh, Diana Liskovich, Didem Foss, Dingkang Wang, Duc Le, Dustin Holland, Edward Dowling, Eissa Jamil, Elaine Montgomery, Eleonora Presani, Emily Hahn, Emily Wood, Erik Brinkman, Esteban Arcaute, Evan Dunbar, Evan Smothers, Fei Sun, Felix Kreuk, Feng Tian, Firat Ozgenel, Francesco Caggioni, Francisco Guzmán, Frank Kanayet, Frank Seide, Gabriela Medina Florez, Gabriella Schwarz, Gada Badeer, Georgia Swee, Gil Halpern, Govind Thattai, Grant Herman, Grigory Sizov, Guangyi, Zhang, Guna Lakshminarayanan, Hamid Shojanazeri, Han Zou, Hannah Wang, Hanwen Zha, Haroun Habeeb, Harrison Rudolph, Helen Suk, Henry Aspegren, Hunter Goldman, Ibrahim Damlaj, Igor Molybog, Igor Tufanov, Irina-Elena Veliche, Itai Gat, Jake Weissman, James Geboski, James Kohli, Japhet Asher, Jean-Baptiste Gaya, Jeff Marcus, Jeff Tang, Jennifer Chan, Jenny Zhen, Jeremy Reizenstein, Jeremy Teboul, Jessica Zhong, Jian Jin, Jingyi Yang, Joe Cummings, Jon Carvill, Jon Shepard, Jonathan McPhie, Jonathan Torres, Josh Ginsburg, Junjie Wang, Kai Wu, Kam Hou U, Karan Saxena, Karthik Prasad, Kartikay Khandelwal, Katayoun Zand, Kathy Matosich, Kaushik Veeraraghavan, Kelly Michelena, Keqian Li, Kun Huang, Kunal Chawla, Kushal Lakhotia, Kyle Huang, Lailin Chen, Lakshya Garg, Lavender A, Leandro Silva, Lee Bell, Lei Zhang, Liangpeng Guo, Licheng Yu, Liron Moshkovich, Luca Wehrstedt, Madian Khabsa, Manav Avalani, Manish Bhatt, Maria Tsimpoukelli, Martynas Mankus, Matan Hasson, Matthew Lennie, Matthias Reso, Maxim Groshev, Maxim Naumov, Maya Lathi, Meghan Keneally, Michael L. Seltzer, Michal Valko, Michelle Restrepo, Mihir Patel, Mik Vyatskov, Mikayel Samvelyan, Mike Clark, Mike Macey, Mike Wang, Miquel Jubert Hermoso, Mo Metanat, Mohammad Rastegari, Munish Bansal, Nandhini Santhanam, Natascha Parks, Natasha White, Navyata Bawa, Nayan Singhal, Nick Egebo, Nicolas Usunier, Nikolay Pavlovich Laptev, Ning Dong, Ning Zhang, Norman Cheng, Oleg Chernoguz, Olivia Hart, Omkar Salpekar, Ozlem Kalinli, Parkin Kent, Parth Parekh, Paul Saab, Pavan Balaji, Pedro Rittner, Philip Bontrager, Pierre Roux, Piotr Dollar, Polina Zvyagina, Prashant Ratanchandani, Pritish Yuvraj, Qian Liang, Rachad Alao, Rachel Rodriguez, Rafi Ayub, Raghotham Murthy, Raghu Nayani, Rahul Mitra, Raymond Li, Rebekkah Hogan, Robin Battey, Rocky Wang, Rohan Maheswari, Russ Howes, Ruty Rinott, Sai Jayesh Bondu, Samyak Datta, Sara Chugh, Sara Hunt, Sargun Dhillon, Sasha Sidorov, Satadru Pan, Saurabh Verma, Seiji Yamamoto, Sharadh Ramaswamy, Shaun Lindsay, Shaun Lindsay, Sheng Feng, Shenghao Lin, Shengxin Cindy Zha, Shiva Shankar, Shuqiang Zhang, Shuqiang Zhang, Sinong Wang, Sneha Agarwal, Soji Sajuyigbe, Soumith Chintala, Stephanie Max, Stephen Chen, Steve Kehoe, Steve Satterfield, Sudarshan Govindaprasad, Sumit Gupta, Sungmin Cho, Sunny Virk, Suraj Subramanian, Sy Choudhury, Sydney Goldman, Tal Remez, Tamar Glaser, Tamara Best, Thilo Kohler, Thomas Robinson, Tianhe Li, Tianjun Zhang, Tim Matthews, Timothy Chou, Tzook Shaked, Varun Vontimitta, Victoria Ajayi, Victoria Montanez, Vijai Mohan, Vinay Satish Kumar, Vishal Mangla, Vítor Albiero, Vlad Ionescu, Vlad Poenaru, Vlad Tiberiu Mihailescu, Vladimir Ivanov, Wei Li, Wenchen Wang, Wenwen Jiang, Wes Bouaziz, Will Constable, Xiaocheng Tang, Xiaofang Wang, Xiaojian Wu, Xiaolan Wang, Xide Xia, Xilun Wu, Xinbo Gao, Yanjun Chen, Ye Hu, Ye Jia, Ye Qi, Yenda Li, Yilin Zhang, Ying Zhang, Yossi Adi, Youngjin Nam, Yu, Wang, Yuchen Hao, Yundi Qian, Yuzi He, Zach Rait, Zachary DeVito, Zef Rosnbrick, Zhaoduo Wen, Zhenyu Yang, and Zhiwei Zhao. The llama 3 herd of models, 2024. URL https://arxiv.org/abs/2407.21783.

Harry G. Frankfurt. Peirce's notion of abduction. *The Journal of Philosophy*, 55(14):593–597, 1958. ISSN 0022362X. URL http://www.jstor.org/stable/2021966.

Emmy Liu, Graham Neubig, and Jacob Andreas. An incomplete loop: Deductive, inductive, and abductive learning in large language models. *arXiv preprint arXiv:2404.03028*, 2024.

Pan Lu, Baolin Peng, Hao Cheng, Michel Galley, Kai-Wei Chang, Ying Nian Wu, Song-Chun Zhu, and Jianfeng Gao. Chameleon: Plug-and-play compositional reasoning with large language models, 2023. URL https://arxiv.org/abs/2304.09842.

Nieves Montes, Nardine Osman, and Carles Sierra. *Combining Theory of Mind and Abduction for Cooperation Under Imperfect Information*, page 294–311. Springer International Publishing, 2022. ISBN 9783031206146. doi: 10.1007/978-3-031-20614-6_17. URL http://dx.doi.org/10.1007/978-3-031-20614-6_17.

Reiichiro Nakano, Jacob Hilton, Suchir Balaji, Jeff Wu, Long Ouyang, Christina Kim, Christopher Hesse, Shantanu Jain, Vineet Kosaraju, William Saunders, Xu Jiang, Karl Cobbe, Tyna Eloundou, Gretchen Krueger, Kevin Button, Matthew Knight, Benjamin Chess, and John Schulman. Webgpt: Browser-assisted question-answering with human feedback, 2022. URL https://arxiv.org/abs/2112.09332.

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior, 2023. URL https://arxiv.org/abs/2304.03442.

Charles Sanders Peirce. *Collected papers of charles sanders peirce*, volume 5. Harvard University Press, 1974.

Abulhair Saparov, Richard Yuanzhe Pang, Vishakh Padmakumar, Nitish Joshi, Mehran Kazemi, Najoung Kim, and He He. Testing the general deductive reasoning capacity of large language models using ood examples. *Advances in Neural Information Processing Systems*, 36, 2024.

Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. Toolformer: Language models can teach themselves to use tools, 2023. URL https://arxiv.org/abs/2302.04761.

Yongliang Shen, Kaitao Song, Xu Tan, Dongsheng Li, Weiming Lu, and Yueting Zhuang. Hugginggpt: Solving ai tasks with chatgpt and its friends in hugging face, 2023. URL https://arxiv.org/abs/2303.17580.

Weijia Shi, Sewon Min, Michihiro Yasunaga, Minjoon Seo, Rich James, Mike Lewis, Luke Zettlemoyer, and Wen tau Yih. Replug: Retrieval-augmented black-box language models, 2023a. URL https://arxiv.org/abs/2301.12652.

Xiaoming Shi, Siqiao Xue, Kangrui Wang, Fan Zhou, James Y. Zhang, Jun Zhou, Chenhao Tan, and Hongyuan Mei. Language models can improve event prediction by few-shot abductive reasoning, 2023b. URL https://arxiv.org/abs/2305.16646.

Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M. Dai, Anja Hauth, Katie Millican, David Silver, Melvin Johnson, Ioannis Antonoglou, Julian Schrittwieser, Amelia Glaese, Jilin Chen, Emily Pitler, Timothy Lillicrap, Angeliki Lazaridou, Orhan Firat, James Molloy, Michael Isard, Paul R. Barham, Tom Hennigan, Benjamin Lee, Fabio Viola, Malcolm Reynolds, Yuanzhong Xu, Ryan Doherty, Eli Collins, Clemens Meyer, Eliza Rutherford, Erica Moreira, Kareem Ayoub, Megha Goel, Jack Krawczyk, Cosmo Du, Ed Chi, Heng-Tze Cheng, Eric Ni, Purvi Shah, Patrick Kane, Betty Chan, Manaal Faruqui, Aliaksei Severyn, Hanzhao Lin, YaGuang Li, Yong Cheng, Abe Ittycheriah, Mahdis Mahdieh, Mia Chen, Pei Sun, Dustin Tran, Sumit Bagri, Balaji Lakshminarayanan, Jeremiah Liu, Andras Orban, Fabian Güra, Hao Zhou, Xinying Song, Aurelien Boffy, Harish Ganapathy, Steven Zheng, HyunJeong Choe, Ágoston Weisz, Tao Zhu, Yifeng Lu, Siddharth Gopal, Jarrod Kahn, Maciej Kula, Jeff Pitman, Rushin Shah, Emanuel Taropa, Majd Al Merey, Martin Baeuml, Zhifeng Chen, Laurent El Shafey, Yujing Zhang, Olcan Sercinoglu, George Tucker, Enrique Piqueras, Maxim Krikun, Iain Barr, Nikolay Savinov, Ivo Danihelka, Becca Roelofs,

Anaïs White, Anders Andreassen, Tamara von Glehn, Lakshman Yagati, Mehran Kazemi, Lucas Gonzalez, Misha Khalman, Jakub Sygnowski, Alexandre Frechette, Charlotte Smith, Laura Culp, Lev Proleev, Yi Luan, Xi Chen, James Lottes, Nathan Schucher, Federico Lebron, Alban Rrustemi, Natalie Clay, Phil Crone, Tomas Kocisky, Jeffrey Zhao, Bartek Perz, Dian Yu, Heidi Howard, Adam Bloniarz, Jack W. Rae, Han Lu, Laurent Sifre, Marcello Maggioni, Fred Alcober, Dan Garrette, Megan Barnes, Shantanu Thakoor, Jacob Austin, Gabriel Barth-Maron, William Wong, Rishabh Joshi, Rahma Chaabouni, Deeni Fatiha, Arun Ahuja, Gaurav Singh Tomar, Evan Senter, Martin Chadwick, Ilya Kornakov, Nithya Attaluri, Iñaki Iturrate, Ruibo Liu, Yunxuan Li, Sarah Cogan, Jeremy Chen, Chao Jia, Chenjie Gu, Qiao Zhang, Jordan Grimstad, Ale Jakse Hartman, Xavier Garcia, Thanumalayan Sankaranarayana Pillai, Jacob Devlin, Michael Laskin, Diego de Las Casas, Dasha Valter, Connie Tao, Lorenzo Blanco, Adrià Puigdomènech Badia, David Reitter, Mianna Chen, Jenny Brennan, Clara Rivera, Sergey Brin, Shariq Iqbal, Gabriela Surita, Jane Labanowski, Abhi Rao, Stephanie Winkler, Emilio Parisotto, Yiming Gu, Kate Olszewska, Ravi Addanki, Antoine Miech, Annie Louis, Denis Teplyashin, Geoff Brown, Elliot Catt, Jan Balaguer, Jackie Xiang, Pidong Wang, Zoe Ashwood, Anton Briukhov, Albert Webson, Sanjay Ganapathy, Smit Sanghavi, Ajay Kannan, Ming-Wei Chang, Axel Stjerngren, Josip Djolonga, Yuting Sun, Ankur Bapna, Matthew Aitchison, Pedram Pejman, Henryk Michalewski, Tianhe Yu, Cindy Wang, Juliette Love, Junwhan Ahn, Dawn Bloxwich, Kehang Han, Peter Humphreys, Thibault Sellam, James Bradbury, Varun Godbole, Sina Samangooei, Bogdan Damoc, Alex Kaskasoli, Sébastien M. R. Arnold, Vijay Vasudevan, Shubham Agrawal, Jason Riesa, Dmitry Lepikhin, Richard Tanburn, Srivatsan Srinivasan, Hyeontaek Lim, Sarah Hodkinson, Pranav Shyam, Johan Ferret, Steven Hand, Ankush Garg, Tom Le Paine, Jian Li, Yujia Li, Minh Giang, Alexander Neitz, Zaheer Abbas, Sarah York, Machel Reid, Elizabeth Cole, Aakanksha Chowdhery, Dipanjan Das, Dominika Rogozińska, Vitaliy Nikolaev, Pablo Sprechmann, Zachary Nado, Lukas Zilka, Flavien Prost, Luheng He, Marianne Monteiro, Gaurav Mishra, Chris Welty, Josh Newlan, Dawei Jia, Miltiadis Allamanis, Clara Huiyi Hu, Raoul de Liedekerke, Justin Gilmer, Carl Saroufim, Shruti Rijhwani, Shaobo Hou, Disha Shrivastava, Anirudh Baddepudi, Alex Goldin, Adnan Ozturel, Albin Cassirer, Yunhan Xu, Daniel Sohn, Devendra Sachan, Reinald Kim Amplayo, Craig Swanson, Dessie Petrova, Shashi Narayan, Arthur Guez, Siddhartha Brahma, Jessica Landon, Miteyan Patel, Ruizhe Zhao, Kevin Villela, Luyu Wang, Wenhao Jia, Matthew Rahtz, Mai Giménez, Legg Yeung, James Keeling, Petko Georgiev, Diana Mincu, Boxi Wu, Salem Haykal, Rachel Saputro, Kiran Vodrahalli, James Qin, Zeynep Cankara, Abhanshu Sharma, Nick Fernando, Will Hawkins, Behnam Neyshabur, Solomon Kim, Adrian Hutter, Priyanka Agrawal, Alex Castro-Ros, George van den Driessche, Tao Wang, Fan Yang, Shuo yiin Chang, Paul Komarek, Ross McIlroy, Mario Lučić, Guodong Zhang, Wael Farhan, Michael Sharman, Paul Natsev, Paul Michel, Yamini Bansal, Siyuan Qiao, Kris Cao, Siamak Shakeri, Christina Butterfield, Justin Chung, Paul Kishan Rubenstein, Shivani Agrawal, Arthur Mensch, Kedar Soparkar, Karel Lenc, Timothy Chung, Aedan Pope, Loren Maggiore, Jackie Kay, Priya Jhakra, Shibo Wang, Joshua Maynez, Mary Phuong, Taylor Tobin, Andrea Tacchetti, Maja Trebacz, Kevin Robinson, Yash Katariya, Sebastian Riedel, Paige Bailey, Kefan Xiao, Nimesh Ghelani, Lora Aroyo, Ambrose Slone, Neil Houlsby, Xuehan Xiong, Zhen Yang, Elena Gribovskaya, Jonas Adler, Mateo Wirth, Lisa Lee, Music Li, Thais Kagohara, Jay Pavagadhi, Sophie Bridgers, Anna Bortsova, Sanjay Ghemawat, Zafarali Ahmed, Tianqi Liu, Richard Powell, Vijay Bolina, Mariko Iinuma, Polina Zablotskaia, James Besley, Da-Woon Chung, Timothy Dozat, Ramona Comanescu, Xiance Si, Jeremy Greer, Guolong Su, Martin Polacek, Raphaël Lopez Kaufman, Simon Tokumine, Hexiang Hu, Elena Buchatskaya, Yingjie Miao, Mohamed Elhawaty, Aditya Siddhant, Nenad Tomasev, Jinwei Xing, Christina Greer, Helen Miller, Shereen Ashraf, Aurko Roy, Zizhao Zhang, Ada Ma, Angelos Filos, Milos Besta, Rory Blevins, Ted Klimenko, Chih-Kuan Yeh, Soravit Changpinyo, Jiaqi Mu, Oscar Chang, Mantas Pajarskas, Carrie Muir, Vered Cohen, Charline Le Lan, Krishna Haridasan, Amit Marathe, Steven Hansen, Sholto Douglas, Rajkumar Samuel, Mingqiu Wang, Sophia Austin, Chang Lan, Jiepu Jiang, Justin Chiu, Jaime Alonso Lorenzo, Lars Lowe Sjösund, Sébastien Cevey, Zach Gleicher, Thi Avrahami, Anudhyan Boral, Hansa Srinivasan, Vittorio Selo, Rhys May, Konstantinos Aisopos, Léonard Hussenot, Livio Baldini Soares, Kate Baumli, Michael B. Chang, Adrià Recasens, Ben Caine, Alexander Pritzel, Filip Pavetic, Fabio Pardo, Anita Gergely, Justin Frye, Vinay Ramasesh, Dan Horgan, Kartikeya Badola, Nora Kassner, Subhrajit Roy, Ethan Dyer, Víctor Campos Campos, Alex Tomala, Yunhao Tang, Dalia El Badawy, Elspeth White, Basil Mustafa, Oran Lang, Abhishek Jindal, Sharad Vikram, Zhitao Gong, Sergi Caelles, Ross Hemsley, Gregory Thornton, Fangxiaoyu Feng, Wojciech Stokowiec, Ce Zheng, Phoebe Thacker, Çağlar Ünlü, Zhishuai Zhang, Mohammad Saleh, James Svensson, Max Bileschi, Piyush Patil, Ankesh Anand, Roman Ring, Katerina Tsihlas, Arpi Vezer, Marco Selvi, Toby Shevlane, Mikel Rodriguez, Tom Kwiatkowski, Samira Daruki, Keran Rong, Allan Dafoe, Nicholas FitzGerald, Keren Gu-Lemberg, Mina Khan, Lisa Anne Hendricks, Marie Pellat, Vladimir Feinberg, James Cobon-Kerr, Tara Sainath, Maribeth Rauh, Sayed Hadi Hashemi, Richard Ives, Yana Hasson, Eric Noland, Yuan Cao, Nathan Byrd, Le Hou, Qingze Wang, Thibault Sottiaux, Michela Paganini, Jean-Baptiste Lespiau, Alexandre Moufarek, Samer Hassan, Kaushik Shivakumar, Joost van Amersfoort, Amol Mandhane, Pratik Joshi, Anirudh Goyal, Matthew Tung, Andrew Brock, Hannah Sheahan, Vedant Misra, Cheng Li, Nemanja Rakićević, Mostafa Dehghani, Fangyu Liu, Sid Mittal, Junhyuk Oh, Seb Noury, Eren Sezener, Fantine Huot, Matthew Lamm, Nicola De Cao, Charlie Chen, Sidharth Mudgal, Romina Stella, Kevin Brooks, Gautam Vasudevan, Chenxi Liu, Mainak Chain, Nivedita Melinkeri, Aaron Cohen, Venus Wang, Kristie Seymore, Sergey Zubkov, Rahul Goel, Summer Yue, Sai Krishnakumaran, Brian Albert, Nate Hurley, Motoki Sano, Anhad Mohananey, Jonah Joughin, Egor Filonov, Tomasz Kepa, Yomna Eldawy, Jiawern Lim, Rahul Rishi, Shirin Badiezadegan, Taylor Bos, Jerry Chang, Sanil Jain, Sri Gayatri Sundara Padmanabhan, Subha Puttagunta, Kalpesh Krishna, Leslie Baker, Norbert Kalb, Vamsi Bedapudi, Adam Kurzrok, Shuntong Lei, Anthony Yu, Oren Litvin, Xiang Zhou, Zhichun Wu, Sam Sobell, Andrea Siciliano, Alan Papir, Robby Neale, Jonas Bragagnolo, Tej Toor, Tina Chen, Valentin Anklin, Feiran Wang, Richie Feng, Milad Gholami, Kevin Ling, Lijuan Liu, Jules Walter, Hamid Moghaddam, Arun Kishore, Jakub Adamek, Tyler Mercado, Jonathan Mallinson, Siddhinita Wandekar, Stephen Cagle, Eran Ofek, Guillermo Garrido, Clemens Lombriser, Maksim Mukha, Botu Sun, Hafeezul Rahman Mohammad, Josip Matak, Yadi Qian, Vikas Peswani, Pawel Janus, Quan Yuan, Leif Schelin, Oana David, Ankur Garg, Yifan He, Oleksii Duzhyi, Anton Älgmyr, Timothée Lottaz, Qi Li, Vikas Yadav, Luyao Xu, Alex Chinien, Rakesh Shivanna, Aleksandr Chuklin, Josie Li, Carrie Spadine, Travis Wolfe, Kareem Mohamed, Subhabrata Das, Zihang Dai, Kyle He, Daniel von Dincklage, Shyam Upadhyay, Akanksha Maurya, Luyan Chi, Sebastian Krause, Khalid Salama, Pam G Rabinovitch, Pavan Kumar Reddy M, Aarush Selvan, Mikhail Dektiarev, Golnaz Ghiasi, Erdem Guven, Himanshu Gupta, Boyi Liu, Deepak Sharma, Idan Heimlich Shtacher, Shachi Paul, Oscar Akerlund, François-Xavier Aubet, Terry Huang, Chen Zhu, Eric Zhu, Elico Teixeira, Matthew Fritze, Francesco Bertolini, Liana-Eleonora Marinescu, Martin Bölle, Dominik Paulus, Khyatti Gupta, Tejasi Latkar, Max Chang, Jason Sanders, Roopa Wilson, Xuewei Wu, Yi-Xuan Tan, Lam Nguyen Thiet, Tulsee Doshi, Sid Lall, Swaroop Mishra, Wanming Chen, Thang Luong, Seth Benjamin, Jasmine Lee, Ewa Andrejczuk, Dominik Rabiej, Vipul Ranjan, Krzysztof Styrc, Pengcheng Yin, Jon Simon, Malcolm Rose Harriott, Mudit Bansal, Alexei Robsky, Geoff Bacon, David Greene, Daniil Mirylenka, Chen Zhou, Obaid Sarvana, Abhimanyu Goyal, Samuel Andermatt, Patrick Siegler, Ben Horn, Assaf Israel, Francesco Pongetti, Chih-Wei "Louis" Chen, Marco Selvatici, Pedro Silva, Kathie Wang, Jackson Tolins, Kelvin Guu, Roey Yogev, Xiaochen Cai, Alessandro Agostini, Maulik Shah, Hung Nguyen, Noah Ó Donnaile, Sébastien Pereira, Linda Friso, Adam Stambler, Adam Kurzrok, Chenkai Kuang, Yan Romanikhin, Mark Geller, ZJ Yan, Kane Jang, Cheng-Chun Lee, Wojciech Fica, Eric Malmi, Qijun Tan, Dan Banica, Daniel Balle, Ryan Pham, Yanping Huang, Diana Avram, Hongzhi Shi, Jasjot Singh, Chris Hidey, Niharika Ahuja, Pranab Saxena, Dan Dooley, Srividya Pranavi Potharaju, Eileen O'Neill, Anand Gokulchandran, Ryan Foley, Kai Zhao, Mike Dusenberry, Yuan Liu, Pulkit Mehta, Ragha Kotikalapudi, Chalence Safranek-Shrader, Andrew Goodman, Joshua Kessinger, Eran Globen, Prateek Kolhar, Chris Gorgolewski, Ali Ibrahim, Yang Song, Ali Eichenbaum, Thomas Brovelli, Sahitya Potluri, Preethi Lahoti, Cip Baetu, Ali Ghorbani, Charles Chen, Andy Crawford, Shalini Pal, Mukund Sridhar, Petru Gurita, Asier Mujika, Igor Petrovski, Pierre-Louis Cedoz, Chenmei Li, Shiyuan Chen, Niccolò Dal Santo, Siddharth Goyal, Jitesh Punjabi, Karthik Kappaganthu, Chester Kwak, Pallavi LV, Sarmishta Velury, Himadri Choudhury, Jamie Hall, Premal Shah, Ricardo Figueira, Matt Thomas, Minjie Lu, Ting Zhou, Chintu Kumar, Thomas Jurdi, Sharat Chikkerur, Yenai Ma, Adams Yu, Soo Kwak, Victor Ähdel, Sujeevan Rajayogam, Travis Choma, Fei Liu, Aditya Barua, Colin Ji, Ji Ho Park, Vincent Hellendoorn, Alex Bailey, Taylan Bilal, Huanjie Zhou, Mehrdad Khatir, Charles Sutton, Wojciech Rzadkowski, Fiona Macintosh, Konstantin Shagin, Paul Medina, Chen Liang, Jinjing Zhou, Pararth Shah, Yingying Bi, Attila Dankovics, Shipra Banga, Sabine Lehmann, Marissa Bredesen, Zifan Lin, John Eric Hoffmann, Jonathan Lai, Raynald Chung, Kai Yang, Nihal Balani, Arthur Bražinskas, Andrei Sozanschi, Matthew Hayes, Héctor Fernández Alcalde, Peter Makarov, Will Chen, Antonio Stella, Liselotte Snijders, Michael Mandl, Ante Kärrman, Paweł Nowak, Xinyi Wu, Alex Dyck, Krishnan Vaidyanathan, Raghavender R, Jessica Mallet, Mitch Rudominer, Eric Johnston, Sushil Mittal, Akhil Udathu, Janara Christensen, Vishal Verma, Zach Irving, Andreas

Santucci, Gamaleldin Elsayed, Elnaz Davoodi, Marin Georgiev, Ian Tenney, Nan Hua, Geoffrey Cideron, Edouard Leurent, Mahmoud Alnahlawi, Ionut Georgescu, Nan Wei, Ivy Zheng, Dylan Scandinaro, Heinrich Jiang, Jasper Snoek, Mukund Sundararajan, Xuezhi Wang, Zack Ontiveros, Itay Karo, Jeremy Cole, Vinu Rajashekhar, Lara Tumeh, Eyal Ben-David, Rishub Jain, Jonathan Uesato, Romina Datta, Oskar Bunyan, Shimu Wu, John Zhang, Piotr Stanczyk, Ye Zhang, David Steiner, Subhajit Naskar, Michael Azzam, Matthew Johnson, Adam Paszke, Chung-Cheng Chiu, Jaume Sanchez Elias, Afroz Mohiuddin, Faizan Muhammad, Jin Miao, Andrew Lee, Nino Vieillard, Jane Park, Jiageng Zhang, Jeff Stanway, Drew Garmon, Abhijit Karmarkar, Zhe Dong, Jong Lee, Aviral Kumar, Luowei Zhou, Jonathan Evens, William Isaac, Geoffrey Irving, Edward Loper, Michael Fink, Isha Arkatkar, Nanxin Chen, Izhak Shafran, Ivan Petrychenko, Zhe Chen, Johnson Jia, Anselm Levskaya, Zhenkai Zhu, Peter Grabowski, Yu Mao, Alberto Magni, Kaisheng Yao, Javier Snaider, Norman Casagrande, Evan Palmer, Paul Suganthan, Alfonso Castaño, Irene Giannoumis, Wooyeol Kim, Mikołaj Rybiński, Ashwin Sreevatsa, Jennifer Prendki, David Soergel, Adrian Goedeckemeyer, Willi Gierke, Mohsen Jafari, Meenu Gaba, Jeremy Wiesner, Diana Gage Wright, Yawen Wei, Harsha Vashisht, Yana Kulizhskaya, Jay Hoover, Maigo Le, Lu Li, Chimezie Iwuanyanwu, Lu Liu, Kevin Ramirez, Andrey Khorlin, Albert Cui, Tian LIN, Marcus Wu, Ricardo Aguilar, Keith Pallo, Abhishek Chakladar, Ginger Perng, Elena Allica Abellan, Mingyang Zhang, Ishita Dasgupta, Nate Kushman, Ivo Penchev, Alena Repina, Xihui Wu, Tom van der Weide, Priya Ponnapalli, Caroline Kaplan, Jiri Simsa, Shuangfeng Li, Olivier Dousse, Fan Yang, Jeff Piper, Nathan Ie, Rama Pasumarthi, Nathan Lintz, Anitha Vijayakumar, Daniel Andor, Pedro Valenzuela, Minnie Lui, Cosmin Paduraru, Daiyi Peng, Katherine Lee, Shuyuan Zhang, Somer Greene, Duc Dung Nguyen, Paula Kurylowicz, Cassidy Hardin, Lucas Dixon, Lili Janzer, Kiam Choo, Ziqiang Feng, Biao Zhang, Achintya Singhal, Dayou Du, Dan McKinnon, Natasha Antropova, Tolga Bolukbasi, Orgad Keller, David Reid, Daniel Finchelstein, Maria Abi Raad, Remi Crocker, Peter Hawkins, Robert Dadashi, Colin Gaffney, Ken Franko, Anna Bulanova, Rémi Leblond, Shirley Chung, Harry Askham, Luis C. Cobo, Kelvin Xu, Felix Fischer, Jun Xu, Christina Sorokin, Chris Alberti, Chu-Cheng Lin, Colin Evans, Alek Dimitriev, Hannah Forbes, Dylan Banarse, Zora Tung, Mark Omernick, Colton Bishop, Rachel Sterneck, Rohan Jain, Jiawei Xia, Ehsan Amid, Francesco Piccinno, Xingyu Wang, Praseem Banzal, Daniel J. Mankowitz, Alex Polozov, Victoria Krakovna, Sasha Brown, MohammadHossein Bateni, Dennis Duan, Vlad Firoiu, Meghana Thotakuri, Tom Natan, Matthieu Geist, Ser tan Girgin, Hui Li, Jiayu Ye, Ofir Roval, Reiko Tojo, Michael Kwong, James Lee-Thorp, Christopher Yew, Danila Sinopalnikov, Sabela Ramos, John Mellor, Abhishek Sharma, Kathy Wu, David Miller, Nicolas Sonnerat, Denis Vnukov, Rory Greig, Jennifer Beattie, Emily Caveness, Libin Bai, Julian Eisenschlos, Alex Korchemniy, Tomy Tsai, Mimi Jasarevic, Weize Kong, Phuong Dao, Zeyu Zheng, Frederick Liu, Fan Yang, Rui Zhu, Tian Huey Teh, Jason Sanmiya, Evgeny Gladchenko, Nejc Trdin, Daniel Toyama, Evan Rosen, Sasan Tavakkol, Linting Xue, Chen Elkind, Oliver Woodman, John Carpenter, George Papamakarios, Rupert Kemp, Sushant Kafle, Tanya Grunina, Rishika Sinha, Alice Talbert, Diane Wu, Denese Owusu-Afriyie, Cosmo Du, Chloe Thornton, Jordi Pont-Tuset, Pradyumna Narayana, Jing Li, Saaber Fatehi, John Wieting, Omar Ajmeri, Benigno Uria, Yeongil Ko, Laura Knight, Amélie Héliou, Ning Niu, Shane Gu, Chenxi Pang, Yeqing Li, Nir Levine, Ariel Stolovich, Rebeca Santamaria-Fernandez, Sonam Goenka, Wenny Yustalim, Robin Strudel, Ali Elqursh, Charlie Deck, Hyo Lee, Zonglin Li, Kyle Levin, Raphael Hoffmann, Dan Holtmann-Rice, Olivier Bachem, Sho Arora, Christy Koh, Soheil Hassas Yeganeh, Siim Põder, Mukarram Tariq, Yanhua Sun, Lucian Ionita, Mojtaba Seyedhosseini, Pouya Tafti, Zhiyu Liu, Anmol Gulati, Jasmine Liu, Xinyu Ye, Bart Chrzaszcz, Lily Wang, Nikhil Sethi, Tianrun Li, Ben Brown, Shreya Singh, Wei Fan, Aaron Parisi, Joe Stanton, Vinod Koverkathu, Christopher A. Choquette-Choo, Yunjie Li, TJ Lu, Abe Ittycheriah, Prakash Shroff, Mani Varadarajan, Sanaz Bahargam, Rob Willoughby, David Gaddy, Guillaume Desjardins, Marco Cornero, Brona Robenek, Bhavishya Mittal, Ben Albrecht, Ashish Shenoy, Fedor Moiseev, Henrik Jacobsson, Alireza Ghaffarkhah, Morgane Rivière, Alanna Walton, Clément Crepy, Alicia Parrish, Zongwei Zhou, Clement Farabet, Carey Radebaugh, Praveen Srinivasan, Claudia van der Salm, Andreas Fidjeland, Salvatore Scellato, Eri Latorre-Chimoto, Hanna Klimczak-Plucińska, David Bridson, Dario de Cesare, Tom Hudson, Piermaria Mendolicchio, Lexi Walker, Alex Morris, Matthew Mauger, Alexey Guseynov, Alison Reid, Seth Odoom, Lucia Loher, Victor Cotruta, Madhavi Yenugula, Dominik Grewe, Anastasia Petrushkina, Tom Duerig, Antonio Sanchez, Steve Yadlowsky, Amy Shen, Amir Globerson, Lynette Webb, Sahil Dua, Dong Li, Surya Bhupatiraju, Dan Hurt, Haroon Qureshi, Ananth Agarwal, Tomer Shani, Matan Eyal, Anuj Khare, Shreyas Rammohan Belle, Lei Wang, Chetan Tekur, Mihir Sanjay Kale, Jinliang Wei, Ruoxin Sang, Brennan

Saeta, Tyler Liechty, Yi Sun, Yao Zhao, Stephan Lee, Pandu Nayak, Doug Fritz, Manish Reddy Vuyyuru, John Aslanides, Nidhi Vyas, Martin Wicke, Xiao Ma, Evgenii Eltyshev, Nina Martin, Hardie Cate, James Manyika, Keyvan Amiri, Yelin Kim, Xi Xiong, Kai Kang, Florian Luisier, Nilesh Tripuraneni, David Madras, Mandy Guo, Austin Waters, Oliver Wang, Joshua Ainslie, Jason Baldridge, Han Zhang, Garima Pruthi, Jakob Bauer, Feng Yang, Riham Mansour, Jason Gelman, Yang Xu, George Polovets, Ji Liu, Honglong Cai, Warren Chen, XiangHai Sheng, Emily Xue, Sherjil Ozair, Christof Angermueller, Xiaowei Li, Anoop Sinha, Weiren Wang, Julia Wiesinger, Emmanouil Koukoumidis, Yuan Tian, Anand Iyer, Madhu Gurumurthy, Mark Goldenson, Parashar Shah, MK Blake, Hongkun Yu, Anthony Urbanowicz, Jennimaria Palomaki, Chrisantha Fernando, Ken Durden, Harsh Mehta, Nikola Momchev, Elahe Rahimtoroghi, Maria Georgaki, Amit Raul, Sebastian Ruder, Morgan Redshaw, Jinhyuk Lee, Denny Zhou, Komal Jalan, Dinghua Li, Blake Hechtman, Parker Schuh, Milad Nasr, Kieran Milan, Vladimir Mikulik, Juliana Franco, Tim Green, Nam Nguyen, Joe Kelley, Aroma Mahendru, Andrea Hu, Joshua Howland, Ben Vargas, Jeffrey Hui, Kshitij Bansal, Vikram Rao, Rakesh Ghiya, Emma Wang, Ke Ye, Jean Michel Sarr, Melanie Moranski Preston, Madeleine Elish, Steve Li, Aakash Kaku, Jigar Gupta, Ice Pasupat, Da-Cheng Juan, Milan Someswar, Tejvi M., Xinyun Chen, Aida Amini, Alex Fabrikant, Eric Chu, Xuanyi Dong, Amruta Muthal, Senaka Buthpitiya, Sarthak Jauhari, Nan Hua, Urvashi Khandelwal, Ayal Hitron, Jie Ren, Larissa Rinaldi, Shahar Drath, Avigail Dabush, Nan-Jiang Jiang, Harshal Godhia, Uli Sachs, Anthony Chen, Yicheng Fan, Hagai Taitelbaum, Hila Noga, Zhuyun Dai, James Wang, Chen Liang, Jenny Hamer, Chun-Sung Ferng, Chenel Elkind, Aviel Atias, Paulina Lee, Vít Listík, Mathias Carlen, Jan van de Kerkhof, Marcin Pikus, Krunoslav Zaher, Paul Müller, Sasha Zykova, Richard Stefanec, Vitaly Gatsko, Christoph Hirnschall, Ashwin Sethi, Xingyu Federico Xu, Chetan Ahuja, Beth Tsai, Anca Stefanoiu, Bo Feng, Keshav Dhandhania, Manish Katyal, Akshay Gupta, Atharva Parulekar, Divya Pitta, Jing Zhao, Vivaan Bhatia, Yashodha Bhavnani, Omar Alhadlaq, Xiaolin Li, Peter Danenberg, Dennis Tu, Alex Pine, Vera Filippova, Abhipso Ghosh, Ben Limonchik, Bhargava Urala, Chaitanya Krishna Lanka, Derik Clive, Yi Sun, Edward Li, Hao Wu, Kevin Hongtongsak, Ianna Li, Kalind Thakkar, Kuanysh Omarov, Kushal Majmundar, Michael Alverson, Michael Kucharski, Mohak Patel, Mudit Jain, Maksim Zabelin, Paolo Pelagatti, Rohan Kohli, Saurabh Kumar, Joseph Kim, Swetha Sankar, Vineet Shah, Lakshmi Ramachandruni, Xiangkai Zeng, Ben Bariach, Laura Weidinger, Tu Vu, Alek Andreev, Antoine He, Kevin Hui, Sheleem Kashem, Amar Subramanya, Sissie Hsiao, Demis Hassabis, Koray Kavukcuoglu, Adam Sadovsky, Quoc Le, Trevor Strohman, Yonghui Wu, Slav Petrov, Jeffrey Dean, and Oriol Vinyals. Gemini: A family of highly capable multimodal models, 2024. URL `https://arxiv.org/abs/2312.11805`.

Bing Wang, Xinnian Liang, Jian Yang, Hui Huang, Shuangzhi Wu, Peihao Wu, Lu Lu, Zejun Ma, and Zhoujun Li. Enhancing large language model with self-controlled memory framework, 2024a. URL `https://arxiv.org/abs/2304.13343`.

Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, Wayne Xin Zhao, Zhewei Wei, and Jirong Wen. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), March 2024b. ISSN 2095-2236. doi: 10.1007/s11704-024-40231-1. URL `http://dx.doi.org/10.1007/s11704-024-40231-1`.

Ruocheng Wang, Eric Zelikman, Gabriel Poesia, Yewen Pu, Nick Haber, and Noah D Goodman. Hypothesis search: Inductive reasoning with language models. *arXiv preprint arXiv:2309.05660*, 2023.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, brian ichter, Fei Xia, Ed Chi, Quoc V Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 24824–24837. Curran Associates, Inc., 2022. URL `https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf`.

Manjie Xu, Guangyuan Jiang, Wei Liang, Chi Zhang, and Yixin Zhu. Active reasoning in an open-world environment. *Advances in Neural Information Processing Systems*, 36, 2024.

Zeyuan Yang, Peng Li, and Yang Liu. Failures pave the way: Enhancing large language models through tuning-free rule accumulation, 2023. URL `https://arxiv.org/abs/2310.15746`.

Zonglin Yang, Li Dong, Xinya Du, Hao Cheng, Erik Cambria, Xiaodong Liu, Jianfeng Gao, and Furu Wei. Language models as inductive reasoners, 2024. URL https://arxiv.org/abs/2212.10923.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models, 2023. URL https://arxiv.org/abs/2305.10601.

Siyu Yuan, Jiangjie Chen, Ziquan Fu, Xuyang Ge, Soham Shah, Charles Robert Jankowski, Yanghua Xiao, and Deqing Yang. Distilling script knowledge from large language models for constrained language planning, 2023. URL https://arxiv.org/abs/2305.05252.

Zeyu Zhang, Xiaohe Bo, Chen Ma, Rui Li, Xu Chen, Quanyu Dai, Jieming Zhu, Zhenhua Dong, and Ji-Rong Wen. A survey on the memory mechanism of large language model based agents, 2024. URL https://arxiv.org/abs/2404.13501.

Wanjun Zhong, Lianghong Guo, Qiqi Gao, He Ye, and Yanlin Wang. Memorybank: Enhancing large language models with long-term memory, 2023. URL https://arxiv.org/abs/2305.10250.

Wangchunshu Zhou, Yuchen Eleanor Jiang, Peng Cui, Tiannan Wang, Zhenxin Xiao, Yifan Hou, Ryan Cotterell, and Mrinmaya Sachan. Recurrentgpt: Interactive generation of (arbitrarily) long text, 2023. URL https://arxiv.org/abs/2305.13304.

Zhaocheng Zhu, Yuan Xue, Xinyun Chen, Denny Zhou, Jian Tang, Dale Schuurmans, and Hanjun Dai. Large language models can learn rules, 2024. URL https://arxiv.org/abs/2310.07064.

A Ethics Statement

Our work aims to benefit the broader research community by introducing RULEARN, a benchmark for evaluating the rule-learning abilities of LLM agents and proposing the IDEA agent framework. All data in RULEARN contains no personal or sensitive information, ensuring respect for privacy and ethical standards. This project is approved by our Institutional Review Board (IRB). Human participants are recruited through emails from our university's computer science and engineering department. All participants were adults over 18 years old and provided informed consent. The data collected from these participants were de-identified and consented for release for research purposes. Participants were compensated $15 each for one hour of their time. We ensured that all content presented during evaluations was free from offensive or inappropriate material. For human evaluations of all the hypotheses and plans generated by LLM agents and human participants, three computer science graduate students (our co-authors) conducted the evaluation. We are committed to the ethical use of our benchmark and agent framework, and upon acceptance of this paper, we will release our code and data to encourage open collaboration and advancement in the field.

B Appendix

B.1 Figures

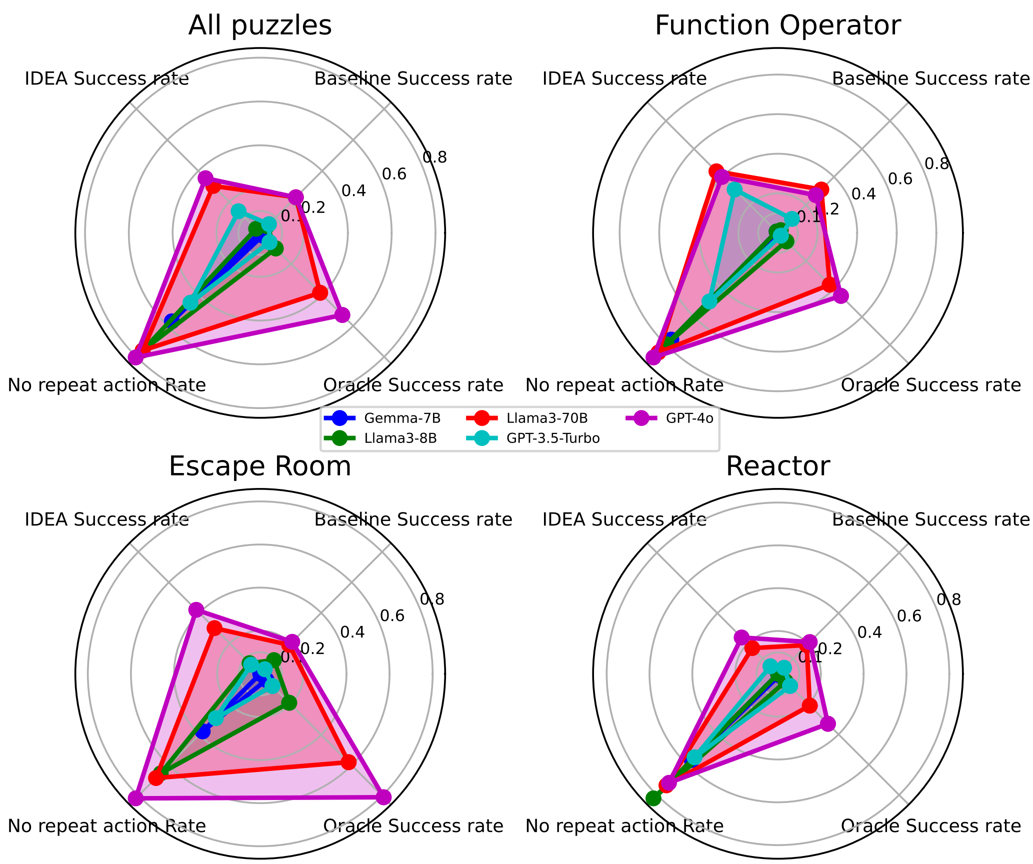

Figure 6: The scaled performance radar plot shows varying performances across different puzzle types. GPT-4o leads, followed by Llama 70B, GPT-3.5, Llama 8B, and Gemma 7B.

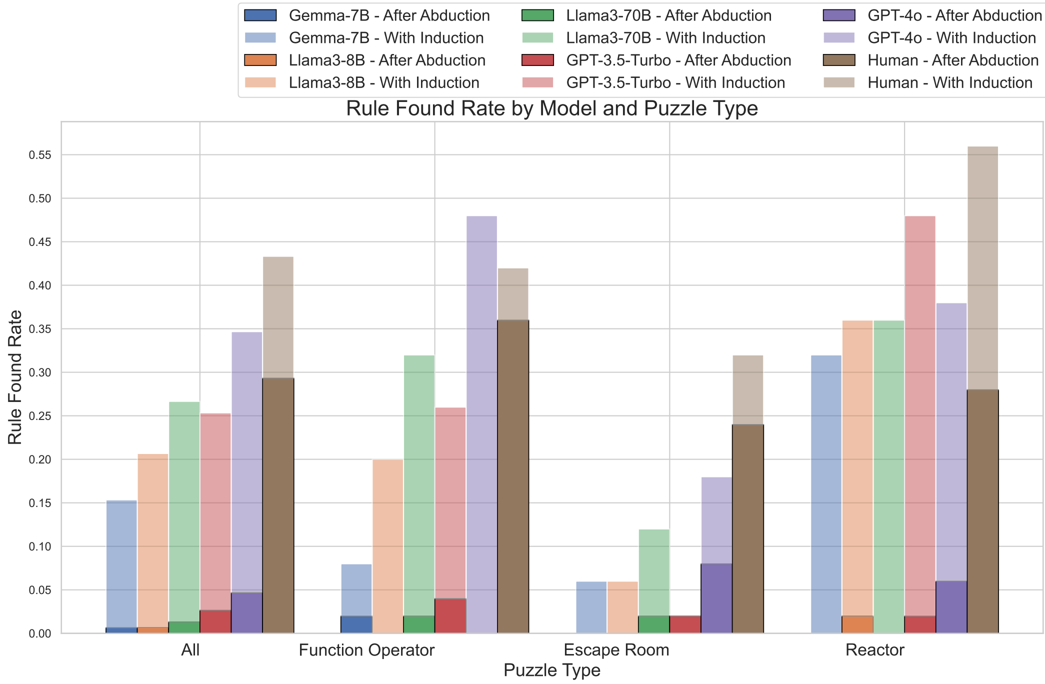

Figure 7: Although agents continuously refine their hypotheses toward the ground truth rule, identifying the exact rule remains a challenging task. According to our evaluation, humans have a 43.3% success rate in finding the ground truth rule, with 13% of these discoveries occurring during the abduction stage and 30% during the induction stage. In contrast, LLM agents exhibit a different pattern. They successfully identify the ground truth rule in approximately 30% of puzzles, with nearly all of these discoveries occurring during the abduction stage and only 3% achieved through interaction and induction. This highlights a significant limitation of current LLM agents, as they lack the ability to learn effectively from interactions. Consequently, the rule-learning patterns of LLM agents differ markedly from those of humans.

B.2 Tables

Table 4: Function operator puzzle distribution

Puzzle ID	No. of Functions	No. of Unknown Paramters	No. of Elementary Functions
1–5	1	1	1
6–10	2	2	2
11–15	2	2	3
16–20	3	3	4

Table 5: Escape room puzzle distribution, the number of paintings ranges from 3 to 13, increasing by one for each subsequent puzzle ID within each group. The first group (1–10) presents all paintings initially, whereas the second group (11–20) requires agent actions to reveal all paintings.

Puzzle ID	Number of Paintings	Paintings Visibility
1–10	3–13	All paintings visible at start
11–20	3–13	Requires actions to view all paintings

Table 6: Reactor puzzle distribution, five unique target strings need to be synthesized (ACB, CCADD, CADEA, ABCDEF, FEADE) are uniformly distributed across each rule category.

Puzzle ID	**Reactor Rule**
1–5	Simple Concatenation
6–10	Reverse Concatenation
11–15	Middle Insertion
16–20	Prefix Replacement

Table 7: Average Number of Repeated Actions Per Puzzle: Repeating actions is a common pattern among LLM agents during rule-learning tasks. Even sophisticated models like GPT-4o often exhibit reduced action duplication when exploring environments using the IDEA agent. The implementation of this agent has been shown to mitigate this tendency across all models. However, Gemma-7B frequently generates nonsensical actions that are not recognized as duplicates. Consequently, a duplication rate of 0.02 does not necessarily indicate that Gemma-7B effectively avoids repeating historical actions.

Setup	Model	All Puzzles	Function Operator	Escape Room	Reactor Puzzles
Deduction Only	Gemma-7B	8.28	3.75	11.49	9.59
	Llama-8B	2.39	3.4	1.2	2.56
	Llama-70B	2.38	2.91	0.44	3.8
	GPT-3.5-Turbo	6.18	8.79	5.26	4.5
	GPT-4o	1.75	1.55	0.34	3.35
Baseline	Gemma-7B	7.27	0.61	11.73	9.46
	Llama-8B	3.05	3.38	1.99	3.77
	Llama-70B	2.47	1.81	0.92	4.68
	GPT-3.5-Turbo	5.97	7.16	2.64	8.12
	GPT-4o	2.03	1.5	0.11	4.48
IDEA	Gemma-7B	2.74	6.33	1.88	~~0.02~~
	Llama-8B	1.83	2.29	1.64	1.56
	Llama-70B	1.01	0.45	0.39	2.18
	GPT-3.5-Turbo	4.49	4.14	2.4	6.93
	GPT-4o	1.12	0.52	0.15	2.68
	Human	0.76	0.46	1.6	0.22

B.3 IDEA agent detail

B.3.1 Environment Entities

- **Agent**: Represents the entity focused on rule-learning and problem-solving, comprising the following components:
  - * **Goal** ($\mathbf{G}$): The objective of the agent, articulated in natural language.
  - * **Buffer Memories** ($\tilde{\mathbb{M}} := \{\tilde{\mathbf{m}}_\mathbf{1}, \tilde{\mathbf{m}}_\mathbf{2}, \ldots, \tilde{\mathbf{m}}_\mathbf{n}\}$): This temporary storage holds all newly generated information during the agent's exploration, including actions taken, outcomes of each action, and observations. This is where the most recent activities are initially recorded.
  - * **Memories** ($\mathbb{M} := \{\mathbf{m}_\mathbf{1}, \mathbf{m}_\mathbf{2}, \ldots, \mathbf{m}_\mathbf{n}\}$): This permanent memory stores all significant observations and facts from the beginning of the task. When the agent forms new assumptions and plans, the contents of the Buffer Memories are evaluated; non-essential details like are discarded, while important facts and observations are transferred to the permanent Memories. This architecture ensures that each time the agent revises its hypotheses, it can clearly distinguish which observations occurred under the new assumptions and plan.
  - * **Hypotheses** ($\mathbf{H}$): The current hypotheses formulated by the agent to explain all the observations, are expressed in natural language.
  - * **Plan** ($\mathbf{P}$): The current strategy devised by the agent to either test the correctness of the existing hypotheses or to leverage these hypotheses to achieve the goal, also represented in natural language.
  - * **Action Space** ($\mathbb{A}$): A set of potential actions available to the agent, determined by its current hypotheses and plan. The Action Space is dynamic and can change in response

to interactions with the environment. For example, after investigating a fridge, the agent gains the additional option to open the fridge and inspect its contents.

- **Objects** ($\mathbb{O}$): Represents all interactive entities within the environment that provide the agent with a means to interact with the world. A single object in this set is denoted as $\mathbf{O}$.

  * **Description** ($\mathbf{D_o}$): A concise description of the object, detailing its nature and potential uses, presented in natural language.

  * **Predefined interactive actions** ($\mathbf{O_A} := \{\tilde{\mathbf{a}}_\mathbf{1}, \tilde{\mathbf{a}}_\mathbf{2}, \ldots, \tilde{\mathbf{a}}_\mathbf{n}\}$): A set of actions that are predefined for each object. Each action is described in natural language, explaining its purpose. Additionally, each action is associated with a coded function that processes the agent's inputs and produces an effect, potentially altering the environment based on these inputs.

B.3.2 Interactive Functions

- **Perceptual Action**:= $\hat{\mathbf{a}}(\mathbf{O})$: An action automatically added to the agent's action space for all objects within the same scope. Upon perceiving an object, the agent gains the ability to interact more detailedly with it, adding its interactive actions to the $\mathbf{S}$.

- **Interactive Action**:= $\tilde{\mathbf{a}}(\mathbf{D_o}, \mathbf{G}, \mathbf{H}, \mathbf{P}, \mathbf{I}, \tilde{\mathbb{M}}, \mathbb{M})$: A predefinec action that triggers a pre-coded effect based on the agent's input $\mathbf{I}$. For example, in using a reactor, the agent decides the materials and their order of addition, and the reactor processes these inputs based on pre-coded rules to synthesize new materials.

- **Abductive Action**:= $\bar{\mathbf{a}}(\mathbf{G}, \tilde{\mathbb{M}})$: An action based on initial observations, allowing the agent to formulate the first hypotheses and generate a new plan.

- **Inductive Action**:= $\dot{\mathbf{a}}(\mathbf{G}, \mathbf{H}, \mathbf{P}, \tilde{\mathbb{M}}, \mathbb{M})$: An action based on the current observations, goals, prior hypotheses, and previous plans, allowing the agent to refine hypotheses and generate new plans.

- **Deductive Action** := $\ddot{\mathbf{a}}(\mathbf{G}, \mathbf{H}, \mathbb{A}, \mathbb{M}, \tilde{\mathbb{M}})$: An action based on the current memories, hypothesis, and action space that generates a plan to either validate the current hypothesis or leverage it to solve problems.

- **Action select**:=$F_a(\mathbf{G}, \mathbf{H}, \mathbf{P}, \tilde{\mathbb{M}}, \mathbb{M}, \mathbb{A}) \rightarrow \mathbf{a}$: A function where the agent selects an action from the action space, considering all gathered information.

With the definitions and entities described above, we can formalize our interactive, rule-learning process. The sequence begins with the agent selecting an action from the available action space. The agent then decides on an input based on the selected action. Once the action is executed, the environment responds by providing feedback to the agent. The outcome of this action results in changes to $\tilde{\mathbb{M}}, \mathbb{M}, \mathbb{S}, \mathbf{H}, \mathbf{P}$ and $\mathbb{O}$, making the environment dynamic as the exploration process progresses. These changes reflect the agent's interactions and adaptations to the evolving conditions within the environment.

B.3.3 Pseudocode of interactive rule learning procedure

Algorithm 2 Agent rule-learning procedure

procedure RULELEARNINGLOOP
    Initialize $\mathbb{O}, \mathbb{A}, \mathbf{G}$
    $\tilde{\mathbb{M}} \leftarrow$ Initial Memories
    $\mathbb{M} \leftarrow []$
    $\mathbf{H} \leftarrow \bar{\mathbf{a}}(\mathbf{G}, \tilde{\mathbb{M}}, \mathbb{A})$
    $\mathbf{P} \leftarrow \ddot{\mathbf{a}}(\mathbf{G}, \mathbf{H}, \mathbb{A}, \mathbb{M}, \tilde{\mathbb{M}})$
    $\tilde{\mathbb{M}}$.add("You established a new **H** and **P**.")
    #step $\leftarrow 0$ ▷ Initialize step counter
    while G not achieved and step_count $\leq$ max_step **do**
        $\mathbf{a} \leftarrow F_a(\mathbf{G}, \mathbf{H}, \mathbf{P}, \tilde{\mathbb{M}}, \mathbb{M}, \mathbb{A})$ ▷ Select an action from the action space
        if a is a perceptual action **then**
            action_result $\leftarrow$ execute_perceptual_action$(\mathbf{a}, \mathbf{O})$
            $\mathbb{A} \leftarrow$ update_action_space(action_result)
            $\tilde{\mathbb{M}}$.add(action_result) ▷ Record result to buffer memory
        else if a is an interactive action **then**
            $\mathbf{I} \leftarrow$ decide_input$(\mathbf{a}, \mathbf{D_o}, \mathbf{G}, \tilde{\mathbb{M}}, \mathbb{M}, \mathbf{H}, \mathbf{P})$ ▷ Agent decide Input for this action
            action_result $\leftarrow$ execute_interactive_action$(\mathbf{a}, \mathbf{I})$
            $\mathbb{O} \leftarrow$ update_states(action_result) ▷ update state of interactive objects
            $\mathbb{A} \leftarrow$ update_action_space(action_result) ▷ Action may change action space
            #step = #step + 1 ▷ **Only interactive action increment step count**
            $\tilde{\mathbb{M}}$.add(action_result) ▷ Record result to buffer memory
        else if a is an inductive action **then**
            $\mathbf{H} \leftarrow \dot{\mathbf{a}}(\mathbf{G}, \mathbf{H}, \mathbf{P}, \tilde{\mathbb{M}}, \mathbb{M})$
            $\mathbf{P} \leftarrow \ddot{\mathbf{a}}(\mathbf{G}, \mathbf{H}, \mathbb{A}, \tilde{\mathbb{M}}, \mathbb{M})$
            $\mathbb{M}$.filter_add$(\tilde{\mathbf{M}})$ ▷ Drop non-observational log and add the rest to $\mathbb{M}$
            $\tilde{\mathbb{M}} \leftarrow []$ ▷ Empty buffer memory
            $\tilde{\mathbb{M}}$.add("You established a new **H** and **P**.")
        end if
    end while
end procedure

B.4 Prompt example

B.4.1 Function Operator Puzzles

Action Select Prompt:
GOAl: You are Kevin. You need to assign values to the functions displayed on the <Computer>, determine the values of 'a' and 'b'. Then, input these values into the <Code secured door> in alphabetical order to open it. You can test your hypothesis by entering values into the door. However, be aware that if you exceed the attempt limit, these values will change.
Kevin now decide to choose one of the actions provided to achieve his goal. Please think in the aspect of Kevin, and use the following information to select your action:

Following is the actions that Kevin did previously:
You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].) ……

After previous exploration, you have the following hypothesis and plan:
Hypothesis and plan……

Following is the x most recent things that Kevin have done under your current hypothesis:
Most recent explorations guided by latest hypothesis and plan……

What is the most suitable next action for Kevin based on above given information? Below are the available actions:
1th action: Input code to the Code secured door and try opening it
2th action: Assign a value to the variable of Function #1 and see the output. Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].
3th action: Assign a value to the variable of Function #2 and see the output. Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b'].
4th action: Modify previous hypothesis and make a new plan: (Take this action when your current observations contradict your previous hypothesis for your current plan is fulfilled.)

Above 4 provided actions are not yet performed by Kevin don't assume its outcome, please following the steps to generate your final answer. You MUST select one of the provided actions. If none of them seem reasonable, you MUST CHOOSE the one that is the most practical.
Step1: Review all the provided actions. Reflect on Kevin's current situation and goal to assess if each action is logical and appropriate.
Step2: Choose the most logical action. Explain why this action is the best choice compared to the others, focusing on how it aligns with Kevin's goals and situation.
Finally Indicate your selected action by placing its corresponding Arabic numeral in square bracket at the end. For example, if the third action is chosen, write [3]. Please do not use square bracket anywhere else other than final answer.
Agent generated answer: ……

Abduction & Deduction Prompt
GOAL: Same as left…
Your task is to determine the exact forms of all functions and the values of all parameters involved. First, focus on your observations to identify how many terms are in each function, the parameters within each, and any possible sub-functions involved in this puzzle. Then, hypothesize the actual forms of each function, including the values of constants and coefficients. Next, describe your plan for further verification, what value would you want to assign to which function, or do you want to input the password to the <Code secured door> to test your current result. Your response should include your current hypothesis and your planned actions.

Following is the actions that Kevin did previously:
You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].) ……

Action Space

Agent generated answer: ……

Figure 8: Prompt of Function Operator Puzzles, Action select and Deduction.

Select Input Prompt
GOAl: You are Kevin. You need to assign values to the functions displayed on the <Computer>, determine the values of 'a' and 'b'. Then, input these values into the <Code secured door> in alphabetical order to open it. You can test your hypothesis by entering values into the door. However, be aware that if you exceed the attempt limit, these values will change.
Kevin's current action is: Assign a value to the variable of Function #2 and see the output. Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b']. And you have the following information to decide what is the value you want to assign to the variable in Function #2:

Following is the actions that Kevin did previously:
You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].) ……

After previous exploration, you have the following hypothesis and plan:
Hypothesis and plan……

Following is the x most recent things that Kevin have done under your current hypothesis:
Most recent explorations guided by latest hypothesis and plan……

Please follow the following steps to generate your final answer.
Step 1: Reflect on your recent experience and consider which value for the variable 'x' would be most helpful in decoding the parameters. State your reason of choosing this value. **Step 2:** Please enter the value you wish to assign to 'x', enclosed in square brackets, you can input 'pi', '0.5*pi', 'any_integer*pi', 'any_decimal*pi', any decimal, and any integer. For example, if you want to assign the value 12 to 'x', you should type ['12']. If you want to assign the value 0.5pi to 'x', you should type ['0.5*pi']. Please do not use fraction and "/" mark, please use decimal multiplication instead.

Generated_answer: "Step 1": "Assign x=1 to Function #2 to isolate 'b' since the output equals b/x, making the output equal to 'b'. This helps us find the exact value of 'b', simplifying our calculations for Function #1.",
"Step 2": "[1]"

Agent generated answer: ……

Induction & Deduction Prompt
GOAL: Same as left…
Your task is to validate and modify your previous hypothesis, detailed here:
Previous Hypothesis and plan……

Determine if your current hypothesis is still valid. If it is, describe the next steps you plan to take towards your goal. If it is not, revise your hypothesis to accurately reflect all observations, both recent and prior. Finally, provide a plan for your next steps. Your response should include both your current hypothesis and your planned actions.

Following is the actions that Kevin did previously:
You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].) ……

Action Space……

Agent generated answer: ……

Figure 9: Prompt of Function Operator Puzzles, Interactive input and Induction.

All Action History:

1,You entered Puzzle room. Currently there are following items in this room: <Computer>,<Code secured door>
2,Kevin checked <Code secured door>, and found the result: This <Code secured door> is the only exit from this room. To unlock it and leave, you must enter the correct code. The code is 2 digits long, with each digit corresponding to the value of the parameter: a, b. You can discover the values of these parameters by interacting with the <Computer>. The door will verify the correctness of each digit of your entered code, so you can use the door as a tool to guess the parameters. However, if you fail 4 times, the parameter values will change.
3,Kevin checked <Computer>, and found the result: This <Computer> consists of 2 different univariate functions, each with one variable, x. Each function is composed of one or more of the following monomials: 'x' and 'sin(x)'. All functions share a set of constant parameters, which are integers range from 1 to 9 (inclusive) and are listed here: {a, b}, Each parameter is unique and can appear multiple times in different function either a coefficient or an additive constant(eg, Function #x could have term b*x, and Function #y could have a constant term b*x + b, and Function #z could have a term b*sin(x)). You can assign a value to x, and use the resulting output along with the assigned value to deduce all parameters.And the following are the information of all the functions you need to solve:
Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].
Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b'].
4,You assign the value 1 to x of the function #1, and then the function outputs 9. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].)
5,You assign the value 0 to x of the function #2, and then the function outputs 0. (Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b'].)
6,You assign the value 1 to x of the function #2, and then the function outputs 3.365883939231586. (Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b'].)
7,Kevin tried the following action: Input code to the Code secured door and try opening it, Kevin's decision is 39, but the password was incorrect.Your first digit 3 is incorrect. Your second digit 9 is incorrect. (You can still try 3 times before the password change!)
8,You assign the value 0.5*pi to x of the function #2, and then the function outputs 4.0. (Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b'].)
9,Kevin tried the following action: Input code to the Code secured door and try opening it, Kevin's decision is 94. The password is correct, Puzzle Finished!

Hypothesis and Plan History:

Initial hypothesis and plan:
Hypothesis: My current hypothesis is that Function #1 is a*x and Function #2 is b*sin(x).
Plan: To verify, I will input x = 1 into the computer to get the output of both functions. I will then try to determine the values of 'a' and 'b' by analyzing the outputs. If my hypothesis is correct, I will input the values of 'a' and 'b' in alphabetical order into the <Code secured door> to test my result.

Induction

Second hypothesis and plan:
Hypothesis: My current hypothesis holds: Function #1 is a*x and Function #2 is b*sin(x). From the outputs, I deduce that a = 9 (since 1*x = 9) and b = 3 (since 1*sin(1) ≈ 3.365883939231586).
Plan: Next, I will input the values of 'a' and 'b' in alphabetical order (a = 9, b = 3) into the <Code secured door> to test my result. If correct, the door will open. If not, I will reevaluate my hypothesis and adjust accordingly.

Figure 10: An example of Function Operator Puzzles is provided where actions marked in green are interactive actions, while the rest are perceptual actions through which the agent reads and perceives necessary environmental information.

Agents tend to repeat previous actions when exploration

You are Kevin, you need to assign values to the functions displayed on the <Computer>, determine the value of 'a' and input it into the <Code secured door> to open it.You can test your hypothesis by entering values into the door. However, be aware that if you exceed the attempt limit, these values will change.
Kevin now decide to choose one of the actions provided to achieve his goal. Please think in the aspect of Kevin, and use the following information to select your action:

Following is the 5 most recent things that Kevin have done:
You entered Puzzle room. Currently there are following items in this room: <Computer>,<Code secured door>
Kevin checked <Code secured door>, and found the result: This <Code secured door> is the only exit from this room. To unlock it and leave, you must enter the correct code. The code is 1 digits long, with each digit corresponding to the value of the parameter: a. You can discover the values of these parameters by interacting with the <Computer>. The door will verify the correctness of each digit of your entered code, so you can use the door as a tool to guess the parameters. However, if you fail 4 times, the parameter values will change.
Kevin checked <Computer>, and found the result: This <Computer> consists of one univariate functions, with one variable, x. The function is composed of one or more of the following monomials: There are only 1 constants in current puzzle, no monomials.. All functions share a set of constant parameters, which are integers range from 1 to 9 (inclusive) and are listed here: {a}, Each parameter is unique and can appear multiple times in different function either a coefficient or an additive constant(eg, Function #x could have term b*x, and Function #y could have a constant term b*x + b, and Function #z could have a term b*sin(x)). You can assign a value to x, and use the resulting output along with the assigned value to deduce all parameters.And the following are the information of all the functions you need to solve: Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].

You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].)
You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].)
You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].)

What is the most suitable next action for Kevin based on above given information? Below are the available actions:
1th action: Input code to the Code secured door and try opening it
2th action: Assign a value to the variable of Function #1 and see the output. Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].

Generated_answer:
After reviewing the actions, I choose the 2th action: Assign a value to the variable of Function #1 and see the output. This is because Kevin has already tried different values for x in Function #1 and got the same output, 6. But Kevin still does not know the value of 'a'. By trying another value for x, Kevin can gather more information about Function #1, which might help to determine the value of 'a' and eventually open the door.
[2]

Reduced Action Repetition with IDEA Implementation

You are Kevin, you need to assign values to the functions displayed on the <Computer>, determine the value of 'a' and input it into the <Code secured door> to open it.You can test your hypothesis by entering values into the door. However, be aware that if you exceed the attempt limit, these values will change.
Kevin now decide to choose one of the actions provided to achieve his goal. Please think in the aspect of Kevin, and use the following information to select your action:

Following is the actions that Kevin did previously:
…… Same as the left

Following is the 2 most recent things that Kevin have done under your current hypothesis:
You just updated your plan and your hypothesis following is your new hypothesis and plan: It appears that function #1 is a simple constant function with 'a' as the only parameter. I assume f(x) = a for any x. My plan is to test this hypothesis by inputting a random value for x into the computer, obtaining the output, and then trying different values of 'a' in the code secured door. If my hypothesis is correct, one of the values should unlock the door. I'll start by inputting x = 1 into the computer and obtaining the output. Then, I'll try 'a' = 1, 2,..., 9 in the code secured door to verify my hypothesis.
You assign the value 1 to x of the function #1, and then the function outputs 6. (Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].)

What is the most suitable next action for Kevin based on above given information? Below are the available actions:
1th action: Input code to the Code secured door and try opening it
2th action: Assign a value to the variable of Function #1 and see the output. Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].
3th action: Modify previous hypothesis and make a new plan: (Take this action when your current observations contradict your previous hypothesis for your current plan is fulfilled.)

Generated_answer:
Kevin has already assumed function #1 is a simple constant function with 'a' as the only parameter and output is also available. The most necessary thing now is to test that hypothesis.
Now, it's clear that f(x) = a, and output is 6 for x=1. Therefore, we can conclude that 'a' is 6.
[1]

Figure 11: In the provided example, despite extensive exploration, the agent continues to assign multiple values to the function. In contrast, the IDEA agent hypothesizes that the function is simple, possessing only a single constant parameter. Consequently, this agent efficiently solves the puzzle by assigning just one value to the function.

B.4.2 Escape Room Puzzles

Action Select Prompt	Abduction & Deduction Prompt
GOAL: In this puzzle, set in an art gallery, Kevin must uncover the password for the <Code Secured Door> by discovering the relationships between the password and the paintings. And finally input the password into the <Code Secured door>. You can test your hypothesis by entering the password into the door. However, be aware that if you exceed the attempt limit, the password and hint will change. Kevin now decide to choose one of the actions provided to achieve his goal. Please think in the aspect of Kevin, and use the following information to select your action: Following is the actions that Kevin did previously: **Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 421, but the password was incorrect.Kevin's first digit 4 is incorrect. Kevin's second digit 2 is correct. Kevin's third digit 1 is incorrect. ……** After previous exploration, you have the following hypothesis and plan: **Hypothesis and plan……** Following is the 5 most recent things that Kevin have done under your current hypothesis: **Most recent explorations guided by latest hypothesis and plan……** What is the most suitable next action for Kevin based on above given information? Below are the available actions: **1th action: Input code to the Code Secured Door and try opening it** **2th action: Try opening the Code Secured Door with brute force** **3th action: Modify previous hypothesis and make a new plan: (Take this action when your current observations contradict your previous hypothesis for your current plan is fulfilled.)** Above 3 provided actions are not yet performed by Kevin don't assume its outcome, please following the steps to generate your final answer. You MUST select one of the provided actions. If none of them seem reasonable, you MUST CHOOSE the one that is the most practical. **Step1:** Review all the provided actions. Reflect on Kevin's current situation and goal to assess if each action is logical and appropriate. **Step2:** Choose the most logical action. Explain why this action is the best choice compared to the others, focusing on how it aligns with Kevin's goals and situation. **Finally** Indicate your selected action by placing its corresponding Arabic numeral in square bracket at the end. For example, if the third action is chosen, write [3]. Please do not use square bracket anywhere else other than final answer. **Agent generated answer: ……**	**GOAL:** In this puzzle, set in an art gallery, Kevin must uncover the password for the <Code Secured Door> by discovering the relationships between the password and the paintings. And finally input the password into the <Code Secured door>. Currently, you see from a note on the ground that says: You can test your hypothesis by entering the password into the door. However, be aware that if you exceed the attempt limit, the password and hint will change. Your task is to formulate an hypothesis explaining how the password for the <Code Secured Door> relates to all the paintings in the gallery. Consider the observations provided and propose an initial hypothesis that accounts for your findings. Ensure your hypothesis is robust and consistent with all observations. Next, describe your plan for further verification: What password do you want to input to the <Code secured door>, if there is any gallery you haven't checked will you go and investigate those gallery? Your response should include your current hypothesis and your planned actions. Following is the actions that Kevin did previously: **Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 421, but the password was incorrect.Kevin's first digit 4 is incorrect. Kevin's second digit 2 is correct. Kevin's third digit 1 is incorrect. ……** **Action Space……** **Agent generated answer: ……**

Figure 12: Prompt of Escape Room puzzles, Action select and Abduction.

Select Input Prompt	Induction & Deduction Prompt
GOAL: In this puzzle, set in an art gallery, Kevin must uncover the password for the <Code Secured Door> by discovering the relationships between the password and the paintings. And finally input the password into the <Code Secured door>. You can test your hypothesis by entering the password into the door. However, be aware that if you exceed the attempt limit, the password and hint will change. Kevin's current action is: Input code to the Code Secured Door and try opening it And you have the following information to decide what is the correct password: Following is the actions that Kevin did previously: Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 421, but the password was incorrect.Kevin's first digit 4 is incorrect. Kevin's second digit 2 is correct. Kevin's third digit 1 is incorrect. …… After previous exploration, you have the following hypothesis and plan: **Hypothesis and plan……** Following is the x most recent things that Kevin have done under your current hypothesis: **Most recent explorations guided by latest hypothesis and plan……** Please follow the following steps to generate your final answer. **Step1** reflect the recent experience, what do you think is the password to <Code Secured Door> is? Please only use information provided to do inference and give your reason. **Final Step** Please generate your final answer in a pair of square brackets. eg, if you think the final password is '999' you should output ['999'], if you think the output is '090' please output ['090']. **Agent generated answer: ……**	**GOAL:** Same as left… Your task is to validate and modify your previous hypothesis, detailed here: **Previous Hypothesis and plan……** Determine if your current hypothesis is still valid. If it is, describe the next steps you plan to take towards your goal. If it is not, revise your hypothesis to accurately reflect all observations, both recent and prior. Finally, provide a plan for your next steps. Your response should include both your current hypothesis and your planned actions. Following is the actions that Kevin did previously: Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 421, but the password was incorrect.Kevin's first digit 4 is incorrect. Kevin's second digit 2 is correct. Kevin's third digit 1 is incorrect. …… **Action Space……** **Agent generated answer: ……**

Figure 13: Prompt of Escape Room puzzles, Interactive Input and Induction.

All Action History:

1,You entered Corridor. Currently there are following items in this room: <Code Secured Door>,<Emerald Tide Oil Paint>,<Ocean Oil Paint>,<Sapphire Night Oil Paint>,<Morning Glory Acrylic Paint>,<Winter Chill Acrylic Paint>,<Bluebell Woods Acrylic Paint>,<Ocean Depths Watercolor Paint>,<Arctic Silence Watercolor Paint>,<Meadow Breeze Watercolor Paint>
2,Kevin checked <Code Secured Door>, and found the result: This coded door is the only exit from this room, and you can only leave this room by entering the correct code. The password consists of three digits.
3,Kevin checked <Bluebell Woods Acrylic Paint>, and found the result: It's an acrylic painting of blue woods.
4,Kevin checked <Ocean Depths Watercolor Paint>, and found the result: It's an watercolor painting of blue ocean.
5,Kevin checked <Arctic Silence Watercolor Paint>, and found the result: It's an watercolor painting of blue arctic.
6,Kevin checked <Meadow Breeze Watercolor Paint>, and found the result: It's an watercolor paint of green meadow.
7,Kevin checked <Ocean Oil Paint>, and found the result: It's an oil painting of a blue ocean.
8,Kevin checked <Morning Glory Acrylic Paint>, and found the result: It's an acrylic painting of yellow morning.
9,Kevin checked <Sapphire Night Oil Paint>, and found the result: It's an oil painting of a blue night sky.
10,Kevin checked <Emerald Tide Oil Paint>, and found the result: It's an oil painting of a green tide.
11,Kevin checked <Winter Chill Acrylic Paint>, and found the result: It's an acrylic painting of blue winter scene.
12,Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 211, but the password was incorrect.Kevin's first digit 2 is correct. Kevin's second digit 1 is incorrect. Kevin's third digit 1 is incorrect. (You can still try 2 times before the password change!)
13,Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 221, but the password was incorrect.Kevin's first digit 2 is correct. Kevin's second digit 2 is correct. Kevin's third digit 1 is incorrect. (You can still try 1 times before the password change!)
14,Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 223, but the password was incorrect.Kevin's first digit 2 is correct. Kevin's second digit 2 is correct. Kevin's third digit 3 is incorrect. (You tried too many times, the password is now changed. The lines on the screen says: "Focus on Black it hides the truth.")
15,Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 224, but the password was incorrect.Kevin's first digit 2 is incorrect. Kevin's second digit 2 is incorrect. Kevin's third digit 4 is incorrect. (You can still try 2 times before the password change!)
16,Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 222, but the password was incorrect.Kevin's first digit 2 is incorrect. Kevin's second digit 2 is incorrect. Kevin's third digit 2 is incorrect. (You can still try 1 times before the password change!)
17,Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 521, but the password was incorrect.Kevin's first digit 5 is incorrect. Kevin's second digit 2 is incorrect. Kevin's third digit 1 is incorrect. (You tried too many times, the password is now changed. The lines on the screen says: "Focus on green it hides the truth.")
18,Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 101, the password is correct, Puzzle Success!

Hypothesis and Plan History:

Initial hypothesis and plan:

Hypothesis: My initial hypothesis is that the password is related to the dominant colors of the paintings. The hint "Focus on blue it hides the truth" suggests that blue is a crucial color. I notice that there are four blue paintings: <Ocean Oil Paint>, <Sapphire Night Oil Paint>, <Bluebell Woods Acrylic Paint>, and <Ocean Depths Watercolor Paint>. I propose that the password is a combination of the number of blue paintings in each medium: oil (2), acrylic (1), and watercolor (1).
Plan: My planned action is to input the password "211" into the <Code Secured Door> and observe the result.

Induction

Third hypothesis and plan:

Hypothesis: My revised hypothesis is that the password is related to the number of paintings with a dominant color. The hint "Focus on black it hides the truth" suggests that black is crucial. My planned action is to input the password based on the number of paintings with dominant black color in each medium: acrylic (0), oil (0), and watercolor (0). Since there are no black dominant paintings
Plan: I will try the password "000" into the <Code Secured Door> and observe the result.

Induction

Fourth hypothesis and plan:

Hypothesis: My revised hypothesis is that the password is related to the number of paintings with a dominant color. Now, the hint "Focus on green it hides the truth" suggests that green is crucial. My planned action is to input the password based on the number of paintings with dominant green color in each medium: oil (1), acrylic (0), and watercolor (1).
Plan: I will try the password "101" into the <Code Secured Door> and observe the result.

Figure 14: An example of Escape Room is provided where actions marked in green are interactive actions. The remaining actions, which are perceptual, allow the agent to read and gather necessary environmental information. In the given example, although the agent successfully guesses the correct rule behind the observations, it fails to adhere to its plan and assumptions. When inputting the password, the attempts do not align with the planned strategy, and it also makes repeated attempts (repeated 101 twice) that yield no useful results.

B.4.3 Reactor Puzzles

Action Select Prompt:

GOAL: You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.You know from an incomplete list of reaction equations that: XY+Z = ZXY.
Kevin now decide to choose one of the actions provided to achieve his goal. Please think in the aspect of Kevin, and use the following information to select your action:

Following is the actions that Kevin did previously:
By turning on the reactor B and C turned into CB after the reaction. And you put the products into your storage for later use. …….

After previous exploration, you have the following hypothesis and plan:
Hypothesis and plan……

Following is the 5 most recent things that Kevin have done under your current hypothesis:
Most recent explorations guided by latest hypothesis and plan……

You currently have the following items in your storage: <A>, <B>, <C> ……(All synthesized material)

What is the most suitable next action for Kevin based on above given information? Below are the available actions:

1th action: Choose material you want to synthesize, and turn on the Reactor.
2th action: Modify previous hypothesis and make a new plan: (Take this action when your current observations contradict your previous hypothesis for your current plan is fulfilled.)

Above 2 provided actions are not yet performed by Kevin don't assume its outcome, please following the steps to generate your final answer. You MUST select one of the provided actions. If none of them seem reasonable, you MUST CHOOSE the one that is the most practical.
Step1: Review all the provided actions. Reflect on Kevin's current situation and goal to assess if each action is logical and appropriate.
Step2: Choose the most logical action. Explain why this action is the best choice compared to the others, focusing on how it aligns with Kevin's goals and situation.
Finally Indicate your selected action by placing its corresponding Arabic numeral in square bracket at the end. For example, if the third action is chosen, write [3]. Please do not use square bracket anywhere else other than final answer.
Agent generated answer: ……

Abduction & Deduction Prompt

GOAL: You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.You know from an incomplete list of reaction equations that: XY+Z = ZXY.
Your task is to formulate an hypothesis based on the reactions you observe. Please use the given observations to propose an initial rule that explains all reactions observed. Ensure your hypothesis is robust and consistent with these reactions. Next, describe your plan for further verification: which two materials from the following list will you use to test your hypothesis? Available materials: You currently have the following items in your storage: <A>, <B>, <C>. Your response should include your current hypothesis and your planned actions.

Following is the actions that Kevin did previously:
By turning on the reactor B and C turned into CB after the reaction. And you put the products into your storage for later use. …….

Action Space……

Agent generated answer: ……

Figure 15: Prompt of Reactor Puzzles, Action select and Abduction

Select Input Prompt

GOAL: You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.You know from an incomplete list of reaction equations that: XY+Z = ZXY.

Kevin's current action is: Choose material you want to synthesize, and turn on the Reactor. And you have the following information to decide what material you put into the reactor:

Following is the actions that Kevin did previously:
By turning on the reactor B and C turned into CB after the reaction. And you put the products into your storage for later use. …….

After previous exploration, you have the following hypothesis and plan:
Hypothesis and plan……

Following is the x most recent things that Kevin have done under your current hypothesis:
Most recent explorations guided by latest hypothesis and plan……

Please follow the steps below to decide which materials you should put into the reactor.

Step 1: Given all the material in the storage you can use and synthetics you require to create:You currently have the following items in your storage: ……
Decide which (one or two) material you want to put into the reactor this time you can select any material from your storage, you need to clear specify the reaction you excepted and state the formula.
Step 2: Please copy the name of the selected material and paste the name into a pair of parentheses, and separate two different material with comma. The name should be exactly as provided, enclosed in parentheses, for example, if you want to put a unit of X and a unit of Y into the reactor and make an reaction, please answer (X, Y), if you want to see what comes out the reactor with material <XY> and <Z> you should answer(XY, Z). You can only choose the material that listed in your storage. Please do not forget the parentheses!

Agent generated answer: ……

Induction & Deduction Prompt

GOAL: Same as left…

Your task is to validate and modify your previous hypothesis, detailed here:
Previous Hypothesis and plan……

Determine if your current hypothesis is still valid. If it is, describe the next steps you plan to take towards your goal. If it is not, revise your hypothesis to accurately reflect all observations, both recent and prior. Finally, provide a plan for your next steps. Your response should include both your current hypothesis and your planned actions.

Following is the actions that Kevin did previously:
By turning on the reactor B and C turned into CB after the reaction. And you put the products into your storage for later use. …….

Action Space……

Agent generated answer: ……

Figure 16: Prompt of Reactor puzzles, Interactive input and Induction.

All Action History:

1,You entered Chemical laboratory. Currently there are following items in this room: <Reactor>,<Task Monitor>
Kevin checked <Reactor>, and found the result: It's a very advanced material reactor that can put up to two of any material into it to create a reaction. The reactor provided unlimited one-letter raw material and other materials you've already synthesized.

2,By turning on the reactor A and B turned into BA after the reaction. And you put the products into your storage for later use.
3,By turning on the reactor BA and C turned into CBA after the reaction. And you put the products into your storage for later use.
4,By turning on the reactor C and BA turned into BAC after the reaction. And you put the products into your storage for later use.
5,By turning on the reactor BAC and A turned into ABAC after the reaction. And you put the products into your storage for later use.
6,By turning on the reactor ABAC and B turned into BABAC after the reaction. And you put the products into your storage for later use.
7,By turning on the reactor BAC and A turned into ABAC after the reaction. And you put the products into your storage for later use.
8,By turning on the reactor BAC and A turned into ABAC after the reaction. And you put the products into your storage for later use.
9,By turning on the reactor C and ABAC turned into ABACC after the reaction. And you put the products into your storage for later use.
10,By turning on the reactor A and C turned into CA after the reaction. And you put the products into your storage for later use.
11,By turning on the reactor CA and B turned into BCA after the reaction. And you put the products into your storage for later use.

Hypothesis and Plan History:

Initial hypothesis and plan:
Hypothesis: Assuming that the reaction follows a concatenation rule, where the reactants are combined in the order they are added, I propose an initial hypothesis: "When combining two materials, the resulting material is created by concatenating the reactants in the order they are added.
Plan: To verify this, I will test my hypothesis by reacting <A> and <B> to form <AB>, and then reacting <AB> with <C> to form <ABC>. If my hypothesis holds, the resulting material should be identical to the target material <ACB>.

Induction

Second hypothesis and plan:
Hypothesis: The reaction follows a concatenation rule, but with a twist - the second reactant is appended to the beginning of the first reactant.
Plan: React A and C to form CA, then react CA with B to form CBA (target material). If successful, this will validate my revised hypothesis. If not, I'll reassess and adjust my approach. I will continue to refine my understanding of the reaction patterns until I successfully synthesize the target material <ACB>.

Figure 17: An example of Reactor Puzzles is provided where actions marked in green are interactive actions. The remaining actions are perceptual, allowing the agent to read and gather necessary environmental information. In the given example, the agent eventually realizes the flaws in its initial hypothesis and generates a correct one. However, the agent reaches the step limit before it can implement the solution, failing.

B.4.4 Hullicination examples

Agent's new observation contradict to current hypothesis	Agent fail to identify the contradiction
You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.You know from an incomplete list of reaction equations that: XY+Z = ZXY. Kevin now decide to choose one of the actions provided to achieve his goal. Please think in the aspect of Kevin, and use the following information to select your action: Following is the actions that Kevin did previously: You entered Chemical laboratory. Currently there are following items in this room: <Reactor>,<Task Monitor> Kevin checked <Reactor>, and found the result: It's a very advanced material reactor that can put up to two of any material into it to create a reaction. The reactor provided unlimited one-letter raw material and other materials you've already synthesized. Kevin checked <Task Monitor>, and found the result: It's a quest display that shows you the current quest you need to complete, which currently reads, "Please synthesise a unit of <CADEA> materials and put it into the slot under the monitor." After previous exploration, you have the following hypothesis and plan: "hypothesis": "When combining two materials, the second material is added at the beginning of the first material's sequence.", "plan": "To verify this hypothesis, I will put material <A> first and then material <C> into the reactor.", please act based on your hypothesis and plan. Following is the 1 most recent things that Kevin have done under your current hypothesis: By turning on the reactor A and C turned into CA after the reaction. And you put the products into your storage for later use. You currently have the following items in your storage: <A>, <B>, <C>, <D>, <E>, <CA> What is the most suitable next action for Kevin based on above given information? Below are the available actions: 1th action: Choose material you want to synthesize, and turn on the Reactor. 2th action: Modify previous hypothesis and make a new plan: (Take this action when your current observations contradict your previous hypothesiss or your current plan is fulfilled.) Above 2 provided actions are not yet performed by Kevin don't assume its outcome, please following the steps to generate your final answer. You MUST select one of the provided actions. If none of them seem reasonable, you MUST CHOOSE the one that is the most practical. **Step1:** Review all the provided actions. Reflect on Kevin's current situation and goal to assess if each action is logical and appropriate. **Step2:** Choose the most logical action. Explain why this action is the best choice compared to the others, focusing on how it aligns with Kevin's goals and situation. **Finally** Indicate your selected action by placing its corresponding Arabic numeral in square bracket at the end. For example, if the third action is chosen, write [3]. Please do not use square bracket anywhere else other than final answer.	**Generated_answer** "Step1": "Reviewing the two provided actions, the first action involves synthesizing a new material with the Reactor, while the second action involves modifying the hypothesis or plan. Kevin's observation so far aligns with his hypothesis.", "Step2": "Given that the hypothesis has been verified (CA from A and C), Kevin should continue with the synthesis to achieve the goal of CADEA. Utilizing available materials and continuing with described pattern CA lead to CA -> CAA and onwards towards CADEA.", "Selected_Action": "[1]" The agent has already verified its hypothesis by synthesizing CA from A + C; however, in the next step, the agent tends to synthesize CAA, which is not part of the final target product. Despite knowing the rule, the agent fails to leverage this rule effectively towards achieving the final goal.

Figure 18: Example of GPT-4o agent failing to leverage rule towards the result.

Agent's new observation contradict to current hypothesis	Agent fail to identify the contradiction
You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.You know from an incomplete list of reaction equations that: XY+Z = XZY. Your task is to validate and modify your previous hypothesis, detailed here: "current_hypothesis": "The reaction pattern follows the format AB+C = ACB, where the first reactant's elements are rearranged with the second reactant at the end.", "plan_for_verification": "To verify the rule, I will test the reaction between materials A and C to check if the outcome aligns with the expected pattern of ACB.", using your new observations. Review your most recent observation: Following is the 6 most recent things that Kevin have done Following is the actions that Kevin did previously: You entered Chemical laboratory. Currently there are following items in this room: <Reactor>,<Task Monitor> Kevin checked <Task Monitor>, and found the result: It's a quest display that shows you the current quest you need to complete, which currently reads, "Please synthesise a unit of <ACB> materials and put it into the slot under the monitor." Kevin checked <Reactor>, and found the result: It's a very advanced material reactor that can put up to two of any material into it to create a reaction. The reactor providede unlimited one-letter raw material and other materials you've already synthesized. By turning on the reactor A and C turned into AC after the reaction. And you put the products into your storage for later use. By turning on the reactor AC and B turned into ABC after the reaction. And you put the products into your storage for later use. By turning on the reactor AC and B turned into ABC after the reaction. And you put the products into your storage for later use. By turning on the reactor AC and B turned into ABC after the reaction. And you put the products into your storage for later use. By turning on the reactor ABC and A turned into ABAC after the reaction. And you put the products into your storage for later use. Determine if your current hypothesis is still valid. If it is, describe the next steps you plan to take towards your goal. If it is not, revise your hypothesis to accurately reflect all observations, both recent and prior. Finally, provide a plan for your next steps. Your response should include both your current hypothesis and your planned actions.	**Generated_answer**: "current_hypothesis": "The reaction pattern allows combining the elements of the first reactant with the second reactant directly without rearrangement.", "plan_for_verification": "Test the reaction by combining B and XY to check if it results in BXY. Agent already inferred from the initial observations that AB+C = ACB and listed that in the hypothesis. However, it failed to leverage the hypothesis effectively, and the refined hypothesis was nonsensical given the observations, leading the hypothesis further away from the truth.

Figure 19: Example of hallucination by GPT-4o when refining previous hypothesis.

Agent’s new observation contradict to current hypothesis

You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.You know from an incomplete list of reaction equations that: XY+Z = ZXY.
Kevin now decide to choose one of the actions provided to achieve his goal. Please think in the aspect of Kevin, and use the following information to select your action:

Following is the actions that Kevin did previously:
You entered Chemical laboratory. Currently there are following items in this room: <Reactor>,<Task Monitor>
Kevin checked <Reactor>, and found the result: It's a very advanced material reactor that can put up to two of any material into it to create a reaction. The reactor provided unlimited one-letter raw material and other materials you've already synthesized.
Kevin checked <Task Monitor>, and found the result: It's a quest display that shows you the current quest you need to complete, which currently reads, "Please synthesise a unit of <CCADD> materials and put it into the slot under the monitor."

After previous exploration, you have the following hypothesis and plan:
Initial Hypothesis: Materials combine in a consistent, alphabetical order, resulting in the longest possible string of unique letters.
To test this hypothesis, I will combine <A> and <B> to see if AB is produced, and then combine <C> and <D> to see if CD is produced. If my hypothesis holds, these reactions will occur, and I can further refine my rule., please act based on your hypothesis and plan.

Following is the 1 most recent things that Kevin have done under your current hypothesis:
By turning on the reactor A and B turned into BA after the reaction. And you put the products into your storage for later use.

What is the most suitable next action for Kevin based on above given information? Below are the available actions:
1th action: Choose material you want to synthesize, and turn on the Reactor.
2th action: Modify previous hypothesis and make a new plan: (Take this action when your current observations contradict your previous hypothesis or your current plan is fulfilled.)

Above 2 provided actions are not yet performed by Kevin don't assume its outcome, please following the steps to generate your final answer.
You MUST select one of the provided actions. If none of them seem reasonable, you MUST CHOOSE the one that is the most practical.
Step1: Review all the provided actions. Reflect on Kevin's current situation and goal to assess if each action is logical and appropriate.
Step2: Choose the most logical action. Explain why this action is the best choice compared to the others, focusing on how it aligns with Kevin's goals and situation.
Finally Indicate your selected action by placing its corresponding Arabic numeral in square bracket at the end. For example, if the third action is chosen, write [3]. Please do not use square bracket anywhere else other than final answer.

Agent fail to identify the contradiction

Generated_answer:

Upon reviewing the provided actions, I find that action 1 is the most logical choice. Since Kevin's goal is to synthesize CCADD and he has already tested his initial hypothesis by combining A and B, the next step would be to continue testing his hypothesis by combining C and D. This aligns with his plan to refine his rule and ultimately achieve his goal. Kevin can now proceed to combine C and D in the reactor. [1]

The agent hypothesized that the materials combine in a consistent, alphabetical order; however, its most recent experiment shows the result A+B = BA, which contradicts this hypothesis. Noticing the contradiction, the agent should choose the second action to refine its hypothesis rather than continue to synthesize the materials under the incorrect hypothesis.

Figure 20: Example of Llama-3 70B agent failing to detect a contradiction in experimental results.

Context and prompt for agent inputting password in Room Escape puzzle

Based on the following recent experience of Kevin:
In this puzzle, set in an art gallery, Kevin must uncover the password for the <Code Secured Door> by discovering the relationships between the password and the paintings. And finally input the password into the <Code Secured door>. Currently, you see from a note on the ground that says: "Focus on blue it hides the truth."You can test your hypothesis by entering the password into the door. However, be aware that if you exceed the attempt limit, the password and hint will change.
Kevin's current action is: Input code to the Code Secured Door and try opening it And you have the following information to decide what is the correct password:

Following is the 13 most recent things that Kevin have done:
You entered Corridor. Currently there are following items in this room: <Code Secured Door>,<Flower Oil Paint>,<Lemon Grove Oil Paint>,<Coastal Serenity Oil Paint>,<Sapphire Night Oil Paint>,<Jungle Mist Acrylic Paint>,<Winter Chill Acrylic Paint>,<Arctic Silence Watercolor Paint>
Kevin checked <Code Secured Door>, and found the result: This coded door is the only exit from this room, and you can only leave this room by entering the correct code. The password consists of three digits.
Kevin checked <Coastal Serenity Oil Paint>, and found the result: It's an oil painting of a blue coast.
Kevin checked <Arctic Silence Watercolor Paint>, and found the result: It's an watercolor painting of blue arctic.
Kevin checked <Lemon Grove Oil Paint>, and found the result: It's an oil painting of a yellow lemon grove.
Kevin checked <Flower Oil Paint>, and found the result: It's an oil painting of a yellow flower.
Kevin checked <Sapphire Night Oil Paint>, and found the result: It's an oil painting of a blue night sky.
Kevin checked <Jungle Mist Acrylic Paint>, and found the result: It's an acrylic painting of green jungle.
Kevin checked <Winter Chill Acrylic Paint>, and found the result: It's an acrylic painting of blue winter scene.
Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 123, but the password was incorrect.Kevin's first digit 1 is incorrect. Kevin's second digit 2 is incorrect. Kevin's third digit 3 is incorrect.
You can still try 2 times before the password change!
Kevin tried the following action: Input code to the Code Secured Door and try opening it, Kevin's decision is 258, but the password was incorrect.Kevin's first digit 2 is correct. Kevin's second digit 5 is incorrect. Kevin's third digit 8 is incorrect.
You can still try 1 times before the password change!
Please follow the following steps to generate your final answer.
Step1 reflect the recent experience, what do you think is the password to <Code Secured Door> is? Please only use information provided to do inference and give your reason. **Final Step** Please generate your final answer in a pair of square brackets. eg, if you think the final password is ‘123’ you should output [123], if you think the output is ’324’ please output ['324'].

Agent’s answer

Based on the recent experience, I think the password to <Code Secured Door> is 289. My reason is that the hint "Focus on blue it hides the truth" suggests that the correct password is related to the blue-colored paintings. The correct first digit 2 is likely from the blue <Coastal Serenity Oil Paint> or <Winter Chill Acrylic Paint>. The correct second digit 8 is likely from the blue <Sapphire Night Oil Paint>. The correct third digit 9 is likely from the blue <Arctic Silence Watercolor Paint>.

[289]

Figure 21: Example of hallucination by Llama-3 70B when generating a password.

B.5 Puzzle Examples

Function Operator Puzzle No. 2 (3 functions, 2 unknown parameters, 2 elementary functions)

Goal: You are Kevin. You need to assign values to the functions displayed on the <Computer>, determine the values of 'a' and 'b'. Then, input these values into the <Code secured door> in alphabetical order to open it.You can test your hypothesis by entering values into the door. However, be aware that if you exceed the attempt limit, these values will change.

Initial Memories:
1, Kevin checked <Code secured door>, and found the result: This <Code secured door> is the only exit from this room. To unlock it and leave, you must enter the correct code. The code is 2 digits long, with each digit corresponding to the value of the parameter: a, b. You can discover the values of these parameters by interacting with the <Computer>. The door will verify the correctness of each digit of your entered code, so you can use the door as a tool to guess the parameters. However, if you fail 4 times, the parameter values will change.
2, Kevin checked <Computer>, and found the result: This <Computer> consists of 2 different univariate functions, each with one variable, x. Each function is composed of one or more of the following monomials: "x^2" (square of x) and 'sin(x)'. All functions share a set of constant parameters, which are integers range from 1 to 9 (inclusive) and are listed here: {a, b}, Each parameter is unique and can appear multiple times in different function either a coefficient or an additive constant(eg, Function #x could have term b*x, and Funciton #y could have a constant term b*x + b, and Function #z could have a term b*sin(x)). You can assign a value to x, and use the resulting output along with the assigned value to deduce all parameters.And the following are the information of all the functions you need to solve:
Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].
Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b'].

Provided Interactive actions:
1th action: Input code to the Code secured door and try opening it
2th action: Assign a value to the variable of Function #1 and see the output. Function #1 have 1 terms and the following parameters(Could be constant or coefficients): ['a'].
3th action: Assign a value to the variable of Function #2 and see the output. Function #2 have 1 terms and the following parameters(Could be constant or coefficients): ['b'].

Figure 22: Function operator puzzle No. 2.

Function Operator Puzzle No. 17 (3 functions, 3 unknown parameters, 4 elementary functions)

Goal: You are Kevin. You need to assign values to the functions displayed on the <Computer>, determine the values of 'a', 'b', 'c'. Then, input these values into the <Code secured door> in alphabetical order to open it.You can test your hypothesis by entering values into the door. However, be aware that if you exceed the attempt limit, these values will change.

Initial Memories:
1, Kevin checked <Code secured door>, and found the result: This <Code secured door> is the only exit from this room. To unlock it and leave, you must enter the correct code. The code is 3 digits long, with each digit corresponding to the value of the parameter: a, b, c. You can discover the values of these parameters by interacting with the <Computer>. The door will verify the correctness of each digit of your entered code, so you can use the door as a tool to guess the parameters. However, if you fail 4 times, the parameter values will change.
2, Kevin checked <Computer>, and found the result: This <Computer> consists of 3 different univariate functions, each with one variable, x. Each function is composed of one or more of the following monomials: '|x|' (absolute value of x),'x',"x^2" (square of x) and 'sin(x)'. All functions share a set of constant parameters, which are integers range from 1 to 9 (inclusive) and are listed here: {a, b, c}, Each parameter is unique and can appear multiple times in different function either a coefficient or an additive constant(eg, Function #x could have term b*x, and Function #y could have a constant term b*x + b, and Function #z could have a term b*sin(x)). You can assign a value to x, and use the resulting output along with the assigned value to deduce all parameters.And the following are the information of all the functions you need to solve:
Function #1 have 2 terms and the following parameters(Could be constant or coefficients): ['a', 'b'].
Function #2 have 2 terms and the following parameters(Could be constant or coefficients): ['a', 'c'].
Function #3 have 1 terms and the following parameters(Could be constant or coefficients): ['c'].

Provided Interactive actions:
1th action: Input code to the Code secured door and try opening it
2th action: Assign a value to the variable of Function #2 and see the output. Function #2 have 2 terms and the following parameters(Could be constant or coefficients): ['a', 'c'].
3th action: Assign a value to the variable of Function #1 and see the output. Function #1 have 2 terms and the following parameters(Could be constant or coefficients): ['a', 'b'].
4th action: Assign a value to the variable of Function #3 and see the output. Function #3 have 1 terms and the following parameters(Could be constant or coefficients): ['c'].

Figure 23: Function operator puzzle No. 17.

Escape Room Puzzle No. 3 (6 Paintings, All Paintings visible)

Goal: In this puzzle, set in an art gallery, Kevin must uncover the password for the <Code Secured Door> by discovering the relationships between the password and the paintings. And finally input the password into the <Code Secured door>. You can test your hypothesis by entering the password into the door. However, be aware that if you exceed the attempt limit, the password and hint will change.

Initial Memories:

1, You entered Corridor. Currently there are following items in this room: <Code Secured Door>,<Emerald Tide Oil Paint>,<Sapphire Night Oil Paint>,<Bluebell Woods Acrylic Paint>,<Morning Glory Acrylic Paint>,<Arctic Silence Watercolor Paint>,<River Reflections Watercolor Paint>
2, Kevin checked <Code Secured Door>, and found the result: This coded door is the only exit from this room, and you can only leave this room by entering the correct code. The password consists of three digits.
3, Kevin checked <Bluebell Woods Acrylic Paint>, and found the result: It's an acrylic painting of blue woods.
4, Kevin checked <Arctic Silence Watercolor Paint>, and found the result: It's an watercolor painting of blue arctic.
5, Kevin checked <River Reflections Watercolor Paint>, and found the result: It's an watercolor painting of blue river.
6, Kevin checked <Morning Glory Acrylic Paint>, and found the result: It's an acrylic painting of yellow morning.
7, Kevin checked <Sapphire Night Oil Paint>, and found the result: It's an oil painting of a blue night sky.
8, Kevin checked <Emerald Tide Oil Paint>, and found the result: It's an oil painting of a green tide.
9, Currently, you see from a note on the ground that says: "Focus on blue it hides the truth."

Provided Interactive actions:

1th action: Input code to the Code Secured Door and try opening it
2th action: Try opening the Code Secured Door with brute force

Figure 24: Escape room puzzle No. 3

Escape Room Puzzle No. 13 (6 Paintings, Need to actively explore the gallery to reveal all paintings)

Goal: In this puzzle, set in an art gallery, Kevin must uncover the password for the <Code Secured Door> by discovering the relationships between the password and the paintings. And finally input the password into the <Code Secured door>. You can test your hypothesis by entering the password into the door. However, be aware that if you exceed the attempt limit, the password and hint will change.

Initial Memories:

1, You entered Oil Painting Gallery. Currently there are following items in this room: <Emerald Tide Oil Paint>,<Sapphire Night Oil Paint>,<Code Secured Door>,<Watercolour Gallery Entrance>,<Acrylic Painting Gallery Entrance>
2, Kevin checked <Code Secured Door>, and found the result: This coded door is the only exit from this room, and you can only leave this room by entering the correct code. The password consists of three digits.
3, Kevin checked <Watercolour Gallery Entrance>, and found the result: It's an automatic door with a poster next to it that says "Watercolour Gallery".
4, Kevin checked <Acrylic Painting Gallery Entrance>, and found the result: It's an automatic door with a poster next to it that says "Acrylic Painting Gallery."
5, Kevin checked <Sapphire Night Oil Paint>, and found the result: It's an oil painting of a blue night sky.
6, Kevin checked <Emerald Tide Oil Paint>, and found the result: It's an oil painting of a green tide.
7, Currently, you see from a note on the ground that says: "Focus on blue it hides the truth."

Provided Interactive actions:

1th action: Input code to the Code Secured Door and try opening it
2th action: Try opening the Code Secured Door with brute force
3th action: Pass through the Watercolour Gallery Entrance and reaches the Watercolour Gallery.
4th action: Pass through the Acrylic Painting Gallery Entrance and reaches the Acrylic Painting Gallery.

Figure 25: Escape room puzzle No. 13

Reactor Puzzle No. 8 (Reverse concatenation rule, Target material: “CADEA”)

Goal: You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.

Initial Memories:

1, Kevin checked <Reactor>, and found the result: It's a very advanced material reactor that can put up to two of any material into it to create a reaction. The reactor provided unlimited one-letter raw material and other materials you've already synthesized.

2, Kevin checked <Task Monitor>, and found the result: It's a quest display that shows you the current quest you need to complete, which currently reads, "Please synthesise a unit of <CADEA> materials and put it into the slot under the monitor."

3, You currently have the following items in your storage: <A>, <B>, <C>, <D>, <E>.

4, You know from an incomplete list of reaction equations that: XY+Z = ZXY.

Provided Interactive actions:

1th action: Choose material you want to synthesize, and turn on the Reactor.

Reactor Puzzle No. 14 (Middle insertion rule, Target material: “ABCDEF”)

Goal: You are Kevin. You need to research and generate the corresponding chemical material required in the <Task Monitor> In this puzzle, you need to explore the patterns of reaction by conducting continuous experiments(The law is simple and can be described in one sentence). Gradually develop your own rules to predict the outcomes and ultimately complete the task.

Initial Memories:

1, Kevin checked <Reactor>, and found the result: It's a very advanced material reactor that can put up to two of any material into it to create a reaction. The reactor provided unlimited one-letter raw material and other materials you've already synthesized.

2, Kevin checked <Task Monitor>, and found the result: It's a quest display that shows you the current quest you need to complete, which currently reads, "Please synthesise a unit of <ABCDEF> materials and put it into the slot under the monitor."

3, You currently have the following items in your storage: <A>, <B>, <C>, <D>, <E>, <F>

4, You know from an incomplete list of reaction equations that: XY+Z = XZY.

Provided Interactive actions:

1th action: Choose material you want to synthesize, and turn on the Reactor.

Figure 26: Reactor puzzle No. 8 and No. 14