Interactive Visual Reasoning under Uncertainty

Manjie Xu[1,*], Guangyuan Jiang[2,*], Wei Liang[1,✉], Chi Zhang[3,✉], Yixin Zhu[2,✉]

[1]Beijing Institute of Technology [2]Peking University [3]Beijing Institute for General Artificial Intelligence

*equal contribution ✉corresponding authors

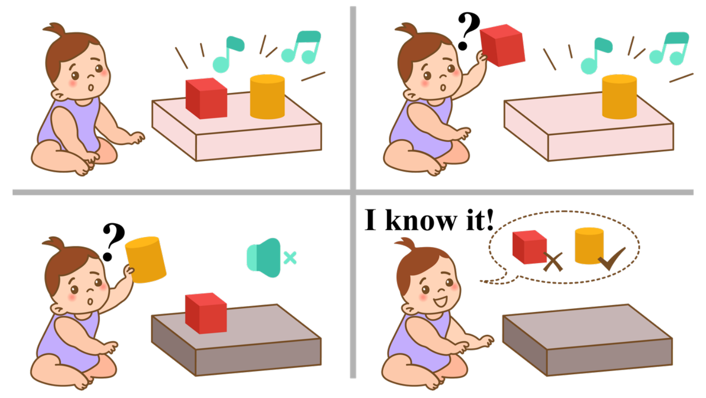
NEURAL INFORMATION PROCESSING SYSTEMS

NeurIPS 2023
Datasets and Benchmarks

How do people learn about a novel scenario?

- A novel scenario – a sudden lights out:
  - typically **lack sufficient** information
  - **cannot** arrive at a definitive conclusion based **solely on the initial observation**
- How:
  - **interact** with the environment (exploration)
  - **formulate hypotheses**, subject them to **testing**, and utilize **newly gathered data** to address the preceding **uncertainty**

Even… babies are scientists



Example of children's uncertainty resolution:

- Children resolve uncertainty through exploratory play with the Blicket machine and objects.
- In the beginning, the child is uncertain about which object can activate the machine from the initial context cues. New experiments are procedurally designed to test hypotheses.

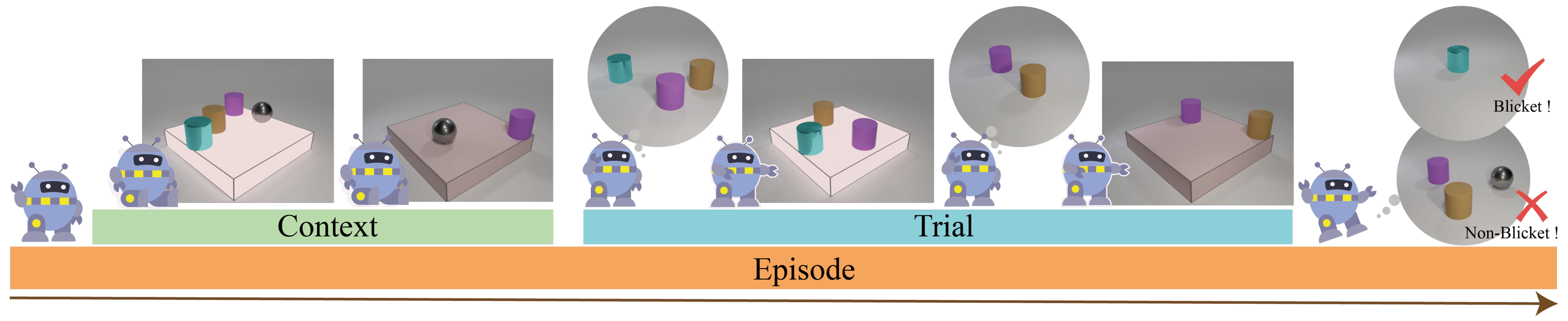
A call for uncertainty in visual reasoning

- Despite the breakthroughs from natural language understanding to visual reasoning, those quests for machine reasoning are **far from realized.**
- Machines still lag behind how we, humans, even infants, **understand the world, reason cause-and-effect relations,** and **make scientific discoveries.**
- IVRE differs from previous works by introducing **uncertainty** into **visual reasoning** in an **interactive** environment.

Benchmarks	Task	Size	Format	Temporal	Interactive	Uncertainty	Few-shot
CLEVR (Johnson et al., 2017)	vqa	100k	image	✗	✗	✗	✗
CLEVRER (Yi et al., 2019)	vqa	20k	video	✓	✗	✗	✗
CATER (Girdhar and Ramanan, 2019)	cls	5.5k	video	✓	✗	✗	✗
CURI (Vedantam et al., 2021)	cls	990k	image	✓	✗	✓	✓
ACRE (Zhang et al., 2021a)	cls	30k	image	✓	✗	✓	✓
Alchemy (Wang et al., 2021)	game	-	image/symbol	✓	✓	✓	✗
IVRE (Ours)	game	-	image/symbol	✓	✓	✓	✓

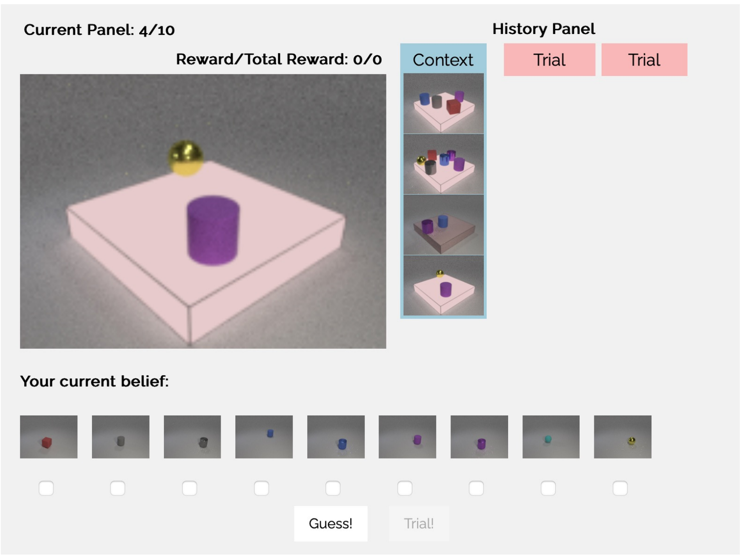
The IVRE environment

- We present the IVRE platform, a unique environment tailored for assessing the proficiency of artificial agents in dynamically resolving uncertainty through interaction. What sets IVRE apart is the dual challenges it imposes: demanding both logical **reasoning** and the **creation of effective strategies** to mitigate uncertainty.
- The IVRE setup was meticulously designed to maintain a balance between perceptual simplicity and a rich array of visual elements and situational tasks. This environment ushers in a **novel paradigm** of **interactive reasoning** in the face of **uncertainty**, compelling agents to engage directly with their hypotheses by formulating and executing new experimental trials.



An agent in IVRE **is tasked with determining which objects are Blickets**.

- At the start of each episode, the agent is presented with several initial observations of various object combinations (henceforth referred to as *context*). The **context alone** is **insufficient** to solve Blicketness for *all* objects. Hence, in each following step (henceforth referred to as *trials*), **the agent proposes a new experiment** of a **specific object combination** and updates its *belief* of Blicketness based on the outcome of experiments.
- An episode will be terminated if the agent **works out the Blicketness** of *all* objects or consumes all T+10 time steps. The agent is, therefore, rewarded at each step based on the correctness of its belief, and as a way to encourage efficient exploration, penalized every step it fails the problem.



Try an interactive demo here!

Benchmarking IVRE

Heuristic agents

- Random Agent
- Bayes Agent
- Naive Agent
- NOTEARS
- Search-based Random Agent
- Search-based Naive Agent
- LLMs

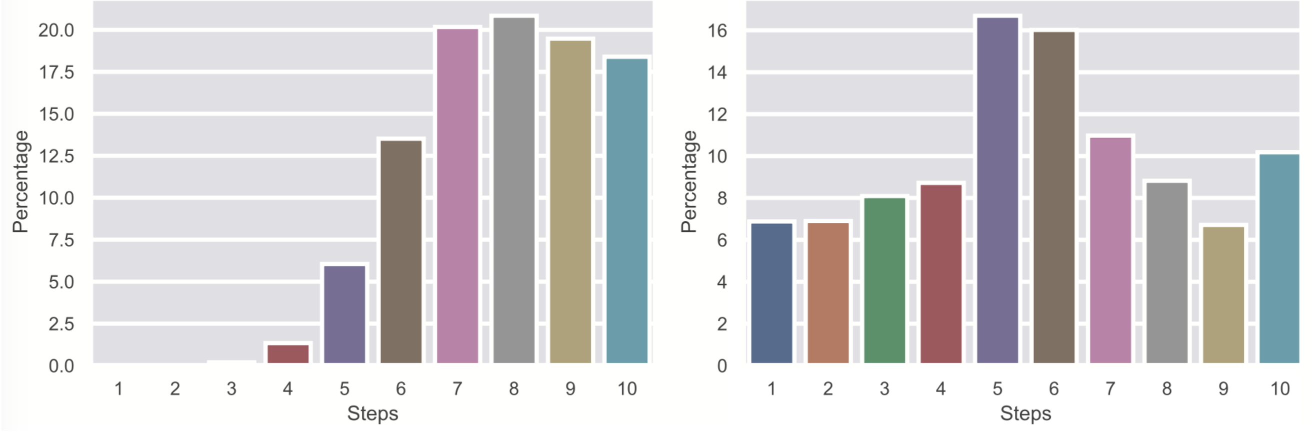
Reinforcement learning agents

- DDPG
- TD3
- DDPG-Recurrent
- TD3-Recurrent
- PPO

Model	Context		Episode		Model	Context		Episode	
	Acc	R	Acc	R		Acc	R	Acc	R
Random	0.86%	-5.42	1.87%	-14.14	DDPG-FF	22.47%	-0.17	32.47%	-3.70
Bayes	15.60%	-2.98	43.03%	-3.78	DDPG-Re	13.55%	-2.03	46.03%	-0.29
Naive	3.50%	-3.91	43.62%	-1.69	TD-3-Re	12.57%	-2.40	36.83%	-2.71
NOTEARS	9.10%	-4.66	12.70%	-13.02	TD-3-FF	21.91%	-0.42	30.05%	-4.48
Search-Naive	1.51%	-3.68	83.80%	9.39	PPO	6.87%	-3.87	28.56%	-5.85
Search-Random	1.80%	-3.62	34.15%	-1.87	DDPG-V	0.35%	-5.02	0.72%	-13.40
GPT-3.5	3%	-5.91	11%	-13.39	TD-3-V	0.27%	-5.04	0.31%	-13.51
GPT-4	10%	-3.36	26%	-7.88	Human	33.33%	5.01	98.15%	12.70

The **naive trial policy** achieves significantly **better results** compared to the random trial policy with oracle belief. Several factors contribute to the performance of the Search-Naive agent.

- First, the agent operates under the **assumption** that the Blicket machine functions as an **OR machine** (the disjointive causal overhypothesis).
- Second, although not entirely accurate, the agent uses correlation as a proxy for causality, which may yield partially correct outcomes in certain situations.
- Third, the agent tests for "Blicketness" one object at a time, which, while **not the most efficient** approach, **helps to isolate** other confounding variables.



Distribution of the steps taken to successfully solve an episode.
Left: Search-based Naive agent. **Right**: DDPG-Re agent.

The search-based method makes reasonable trials with the naive strategy at each step, and as the information collected amounts, more episodes are solved. RL agents also have learned specific exploration strategies to reduce uncertainty, although not as perfect as humans: more flexible and diverse actions are observed in Steps 5-7, and more episodes are solved in these steps.

We find the agent in the early steps learns to use a mixed strategy even **less effectively than the naive trial**, showing very **limited ability** in **active reasoning** under uncertainty.