Sampling Is Not Curiosity

Why LLM Agents Should Investigate

Presenter: Alfonso Amayuelas[1, 2]

With: Piotr Piękos[3], Xin Eric Wang[1], William Yang Wang[1], Jürgen Schmidhuber[3], Alex Pentland[2]

UC Santa Barbara[1], Stanford University[2], KAUST[3]

Curiosity in LLM Agents

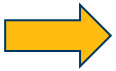
Are LLM agents really curious when they explore?

Curiosity (from RL)

- The agent **acts to learn**, rather than learning while acting
- It keeps a model of the environment and improves it

LLM agents today

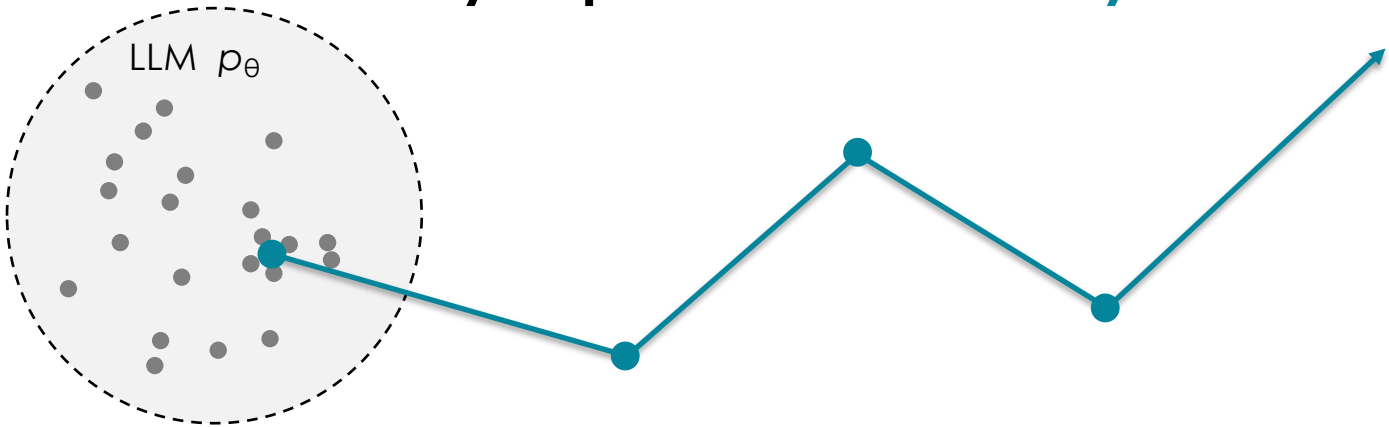
- We call them curious, exploratory, intrinsically motivated
- They explore with temperature, best-of-N, tree search, RAG



These change the sampling or the context,
not the agent's model of the environment

Sampling Is Not Curiosity

Current LLM agents do not explore curiously, they explore stochastically



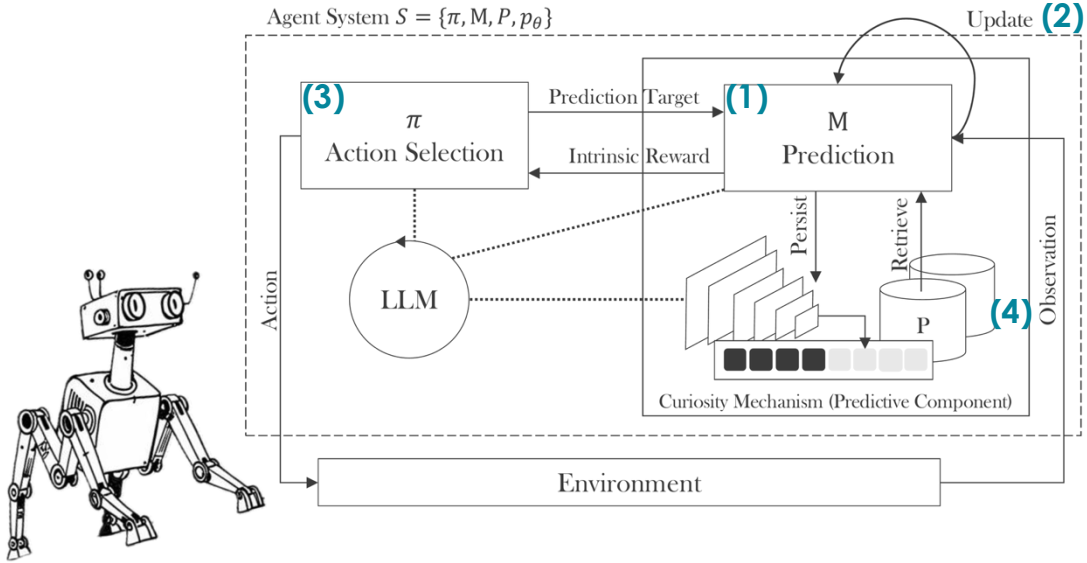
Interaction-level exploration
More samples, the LLM stays fixed

Model-level adaptation
Observations update the agent's model of the environment

Interaction is necessary, but not sufficient for curiosity

Curious Agent Framework

What would a curious agent look like?



(1) **Predictive model (M)**: predicts how the environment responds

(2) **Update**: each observation improves the predictive model

(3) **Action selection (π)**: favors actions the model expects to learn from

(4) **Persistent state (P)**: keeps what was learned beyond the context

The LLM is only one component: curiosity is a property of the whole system

Five Commitments

We reserve "curiosity" for systems that, to some degree,

(1) Predict the environment
how it responds to actions, not their own outputs

(2) Update from interaction
improves within the episode, from what it observes

(3) Act to learn
choose actions expected to improve the prediction

(4) Retain the improvement
beyond the context window

Curious agent
the task is given

(5) Choose what to investigate
set their own prediction targets

Open-ended
it chooses what to predict

Current Practice

Where does current practice fall short?

	Missing
The wrong prediction target The curiosity signal measures the model's own outputs, not the environment	(1)
Memory mistaken for adaptation Stored experience changes behavior, but nothing predicts the environment	(1), (3)
Inquiry direction supplied from outside The agent picks subgoals, but humans define the problem	(5)

Some systems already close the loop for (1) - (4), none for all five

Open Problems

What do we need to build curious agents?

- **Recognizing predictive components**
  Would we recognize a curious system if we saw one?
- **Within-episode measurement**
  Does the agent's prediction actually improve as it acts?
- **Combining with task reward**
  How does the curiosity signal coexist with the agent's actual job?
- **Choosing what to predict**
  What should the agent be trying to predict in the first place?

Takeaway

Sampling is not curiosity

- Today's agents explore stochastically: the sampling and the context change, their model of the environment does not
- Curiosity needs a model of the environment that is **updated**, **drives the actions** and **persists**
- Reserve "curiosity" for systems that meet the commitments, and build toward them

Looking ahead: from completing tasks to open-ended research, aligned with human interests

UC SANTA BARBARA