RETRIEVAL-AUGMENTED LLM AGENTS
Learning to Learn from Experience

NAVER LABS Europe

Thomas Palmeira Ferraz, Romain Deffayet, Vassilina Nikoulina,
Hervé Déjean, **Stéphane Clinchant***

europe.naverlabs.com

ICLR 2026

MemAgents Workshop

Training agents to use retrieved trajectories improves generalization to out-of-distribution tasks

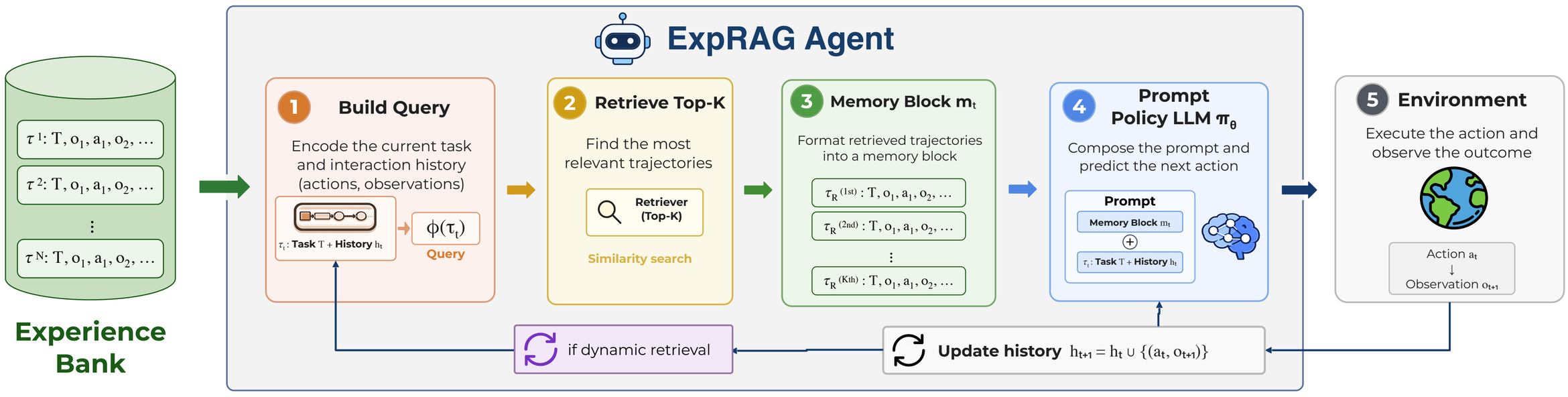
Strong baselines matter
LoRA and retrieval are highly competitive to agent memory and agent training

The Need for stronger Benchmarks
Standard splits hide the real generalization problem

ExpRAG works alone
retrieval already gives large gains by itself

LoRA SFT is not robust alone
Behavior clone can collapse on unseen hard tasks

ExpRAG-LoRA learns to use demonstrations
delivering the strongest OOD generalization



Motivation

- ➔ LLM Agents **struggle to generalize** to unseen tasks.
- ➔ **Behavior Cloning** (Fine-tuning on Expert Trajectories) do well in in-distribution tasks but **collapse out-of-distribution**.
- ➔ Inference-time **agent memory** improves generalization but **underperforms on in-domain scenarios**, as models did not learn completely how to use memories.

Can an LLM-based agent learn to solve new tasks by using retrieved experience in-context?

Strong baselines matter

Simple, well-tuned baselines can outperform more elaborate systems.

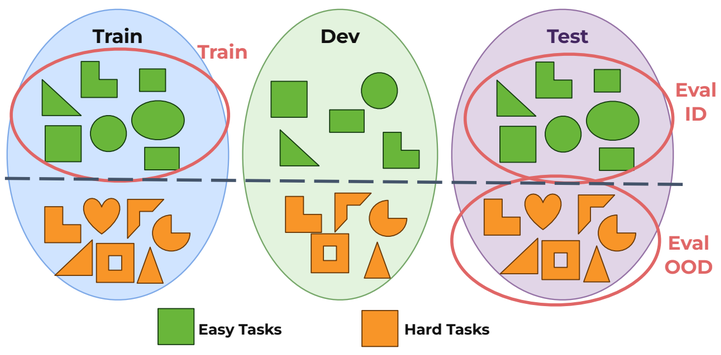
- ➔ Retrieving **raw trajectories beats several complex agent memory** systems
- ➔ Let model train for longer with LoRA SFT beats some **engineered training recipes**, almost **closing benchmark**.

Method	ALFWorld
Prompt-based	
Zero-shot	29.9
ReAct	17.1
Training-Free Memory	
A-MEM	34.7
Reflexion	42.7
Best ExpRAG (ours)	**83.6**
Fine-tuned	
SFT w/ ReAct	80.7
SAND	85.0
Rule-based Expert	89.6
LoRA SFT (on Expert)	**94.1**

Cross-paper comparison on default ALFWorld valid-unseen with Qwen 2.5-7B.

Evaluation protocol: Building OOD Benchmarks

- ➔Benchmarks: Embodied Agents (ALFWorld and ScienceWorld).
- ➔Train trajectories: expert success / Index trajectories: expert (all)
- ➔To evaluate model adaptation at test time, we partition easy/hard from existing splits:



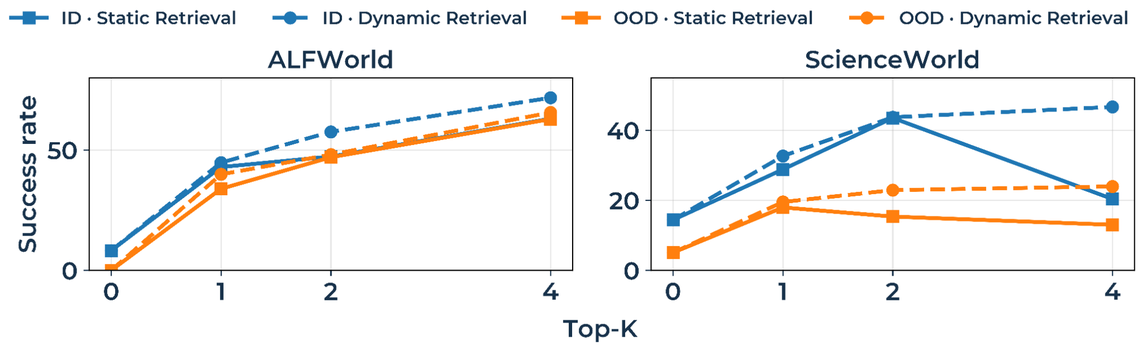
Retrieval Index used

Train / Eval Easy (ID): **train/easy**

Eval Hard (OOD): **train/hard**

Retrieval alone already helps

- ➔ Even without training, retrieval yields large gains over no retrieval.
- ➔ **Top-K**: more trajectories help, until long-context saturation starts to hurt.
- ➔ **Dynamic retrieval**: can improve performance, but introduces instability.



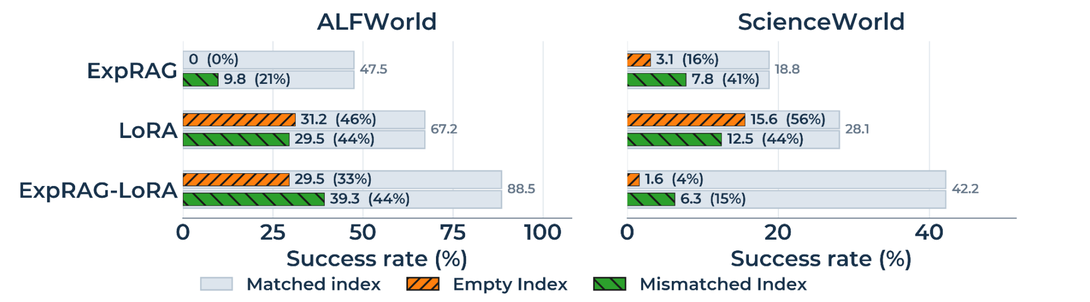
Inference-only ExpRAG Results on Ministral 3-8B

Training with retrieval gives the strongest generalization

Method	ALFWorld		ScienceWorld	
	Easy	Hard	Easy	Hard
Zero-shot	8.2	0.0	14.4	5.1
ExpRAG (inference only)	**54.8**	**47.5**	**43.5**	**18.8**
LoRA (no retrieval)	98.6	34.4	38.8	15.6
LoRA (+ inference ExpRAG)	97.3	67.2	54.1	28.1
ExpRAG-LoRA	**97.3**	**88.5**	**58.8**	**42.2**

Fine-tuning Results on Ministral 3-8B

- ➔ **Fine-tuning alone**: strong in-domain, but **weak on hard unseen tasks**.
- ➔ **Retrieval at inference helps, but training the model to use retrieval helps most. Learn to use trajectories, including failures.**
- ➔ **ExpRAG-LoRA** largely delivers the **strongest OOD generalization**.



Robustness Stress Test - Ministral 3-8B on hard tasks (OOD)

- ➔ **Mismatched:** ExpRAG-LoRA remains **partially robust** on ALFWorld; some gains survive with the easy index on hard tasks.
- ➔ **Empty:** degradation on ALFWorld, but **still competitive to LoRA**.
- ➔ ExpRAG-LoRA **can still be useful** even when demonstrations may not be available for future unseen tasks.

Check out our paper for:

- ➔ **Full results, with other models**
  - Ministral 3-8B, Gemma 3-4B, Qwen 2.5-7B, Qwen 2.5-7B-1M
- ➔ **More analysis** on other **retrieval choices**, **Grokking on LLM Agents**, **Long-context**, **Efficiency Cost**, and more...
- ➔ Practical Recommendations on Training Agents