SelfEvolveAgent: Continual Learning in AI Agents using Memory

Hiva Zaad (Mohammadzadeh)
CS329A: Self-Improving AI Agents, Stanford University

CS329A

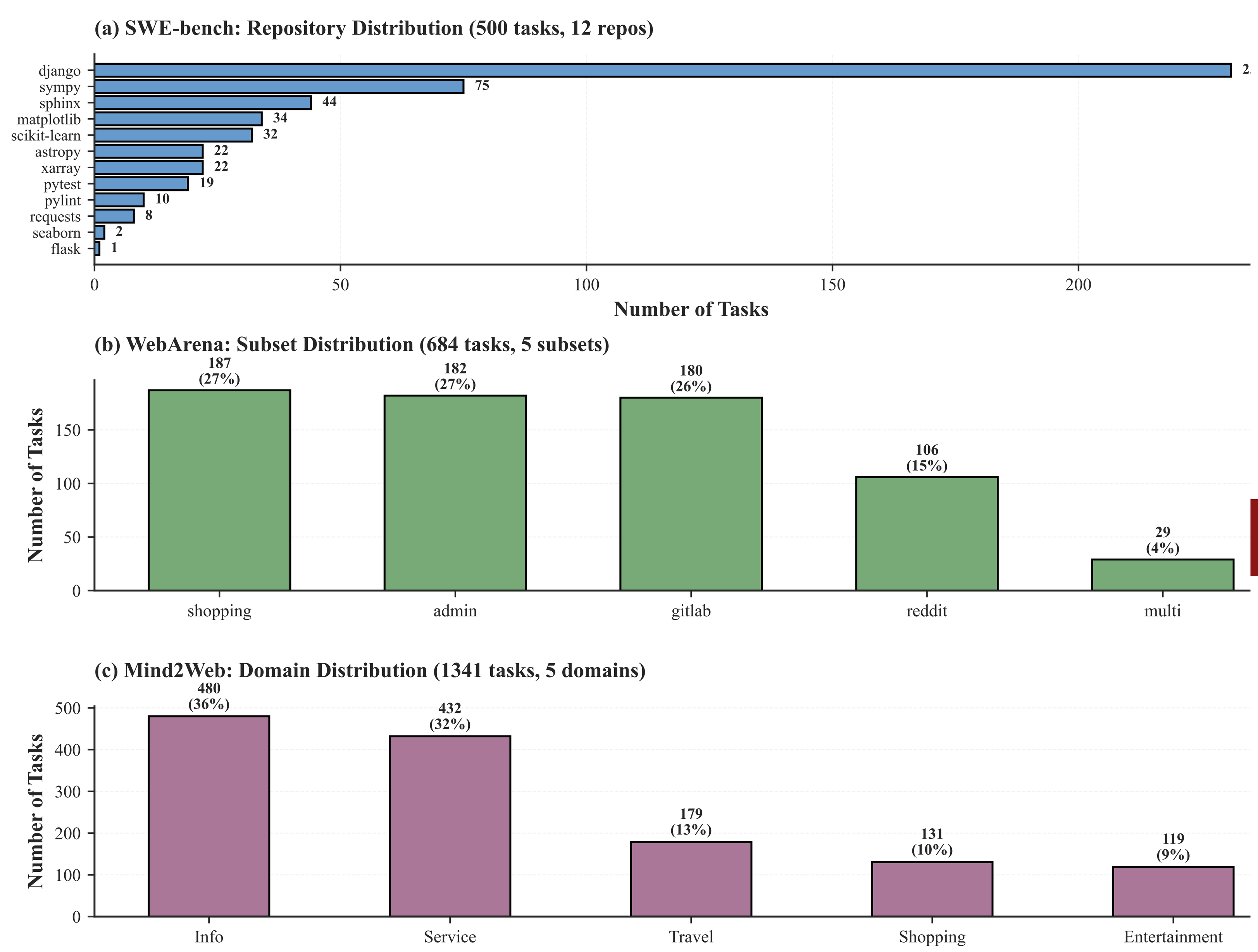
SCAN ME

Introduction and Motivation

- LLM agents excel at long-horizon reasoning but **do not learn** from past experience; often leading to repeated mistakes and weak transfer.
- **Continual learning via memory** is essential for agents that improve over time and adapt to new environments.
- ReasoningBank introduces memory but lacks open-source implementation and analysis of memory quality, difficulty, and retrieval effects.
- **Goal:** Build and analyze SelfEvolveAgent, an open-source ReasoningBank-style agent with rich evaluation of memory quality and utility.

Benchmarks and Setup

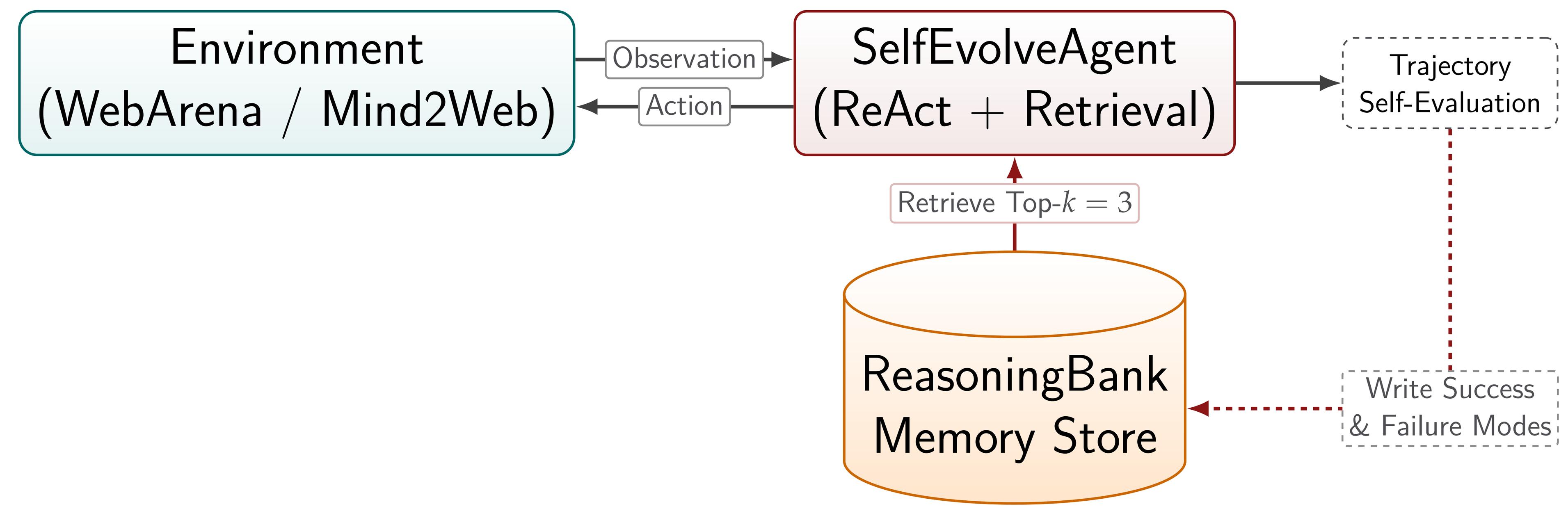
- **Environments**: WebArena, Mind2Web, SWE-Bench.
- **Models**: Gemini-2.5-Flash, Gemini-2.5-Pro, Qwen2.5-72B-Inst
- **Agent Variants**
  - **No Memory**: standard ReAct Agent
  - **ReasoningBank / SelfEvolveAgent**: ReAct agent + memory representation and retrieval.



Cross-benchmark task distributions (SWE-Bench, WebArena, Mind2Web).

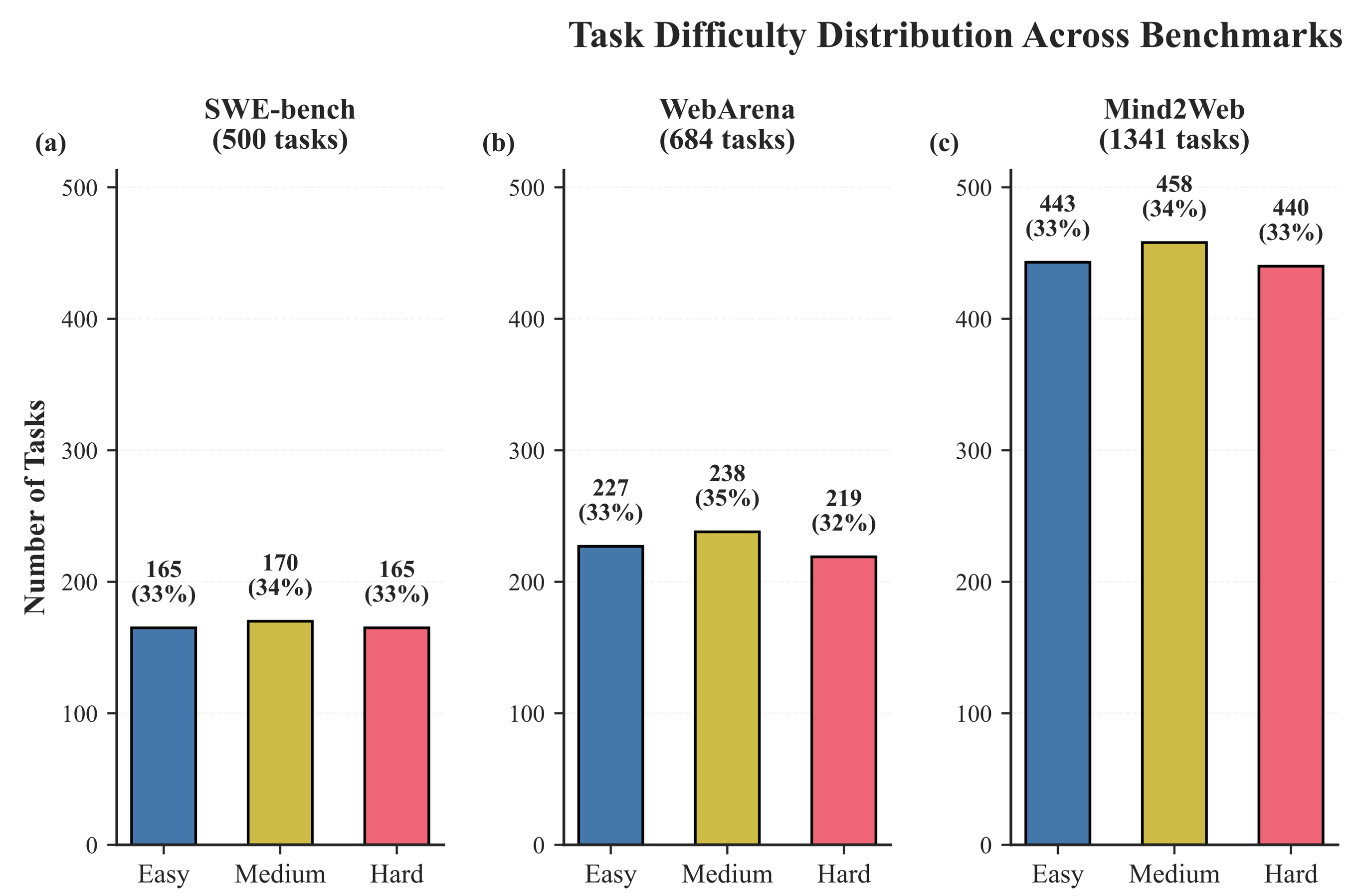
System Overview: From ReAct to SelfEvolveAgent

- No Memory ReAct Think–Act–Observe loop.
- ReasoningBank = ReAct + retrieval + self-evaluation:
  - **Retrieves** top-k relevant strategies (successes & failures).
  - **Updates** memory after self-evaluating the trajectory.
- MemoryOpt: Optimized memory storage and retrieval



Task Difficulty Structure

- Memory yields largest gains on Hard subsets (Admin, Multi, Reddit, SWE-Bench).
- These tasks have long horizons repeated subtask structures.
- Memory can hurt easy tasks (overhead dominates).



Summary Statistics

Benchmark	Easy	Med	Hard	Total
SWE-bench	165 (33%)	170 (34%)	165 (33%)	500
WebArena	227 (33%)	238 (35%)	219 (32%)	684
Mind2Web	443 (33%)	458 (34%)	440 (33%)	1341

Results (Gemini-2.5-Flash)

- **WebArena:**

	No Mem		RBank		MemOpt		NoMem Exp.		RBank Exp.	
Task Subset	SR(%)	Step	SR(%)	Step	SR(%)	Step	SR(%)	Step	SR(%)	Step
Multi (29)	10.34	17.0	13.80	10.0	**17.24**	14.0	10.30	8.8	13.80	8.8
Reddit (106)	68.87	9.62	70.75	10.32	**77.36**	8.1	55.70	6.7	67.00	5.6

- **SWE-Bench-Verified:**

	No Mem		RBank		MemOpt		NoMem Exp.		RBank Exp.	
	SR(%)	Step	SR(%)	Step	SR(%)	Step	SR(%)	Step	SR(%)	Step
Overall	35.30	16.4	38.13	14.15	**39.00**	13.13	34.20	30.3	38.80	27.5

SR = success rate (%), Step = average steps (lower is better).

Memory Representation

- **Memory objects** (JSON):
  - title (strategy), description (summary), content (multi-step reasoning)
  - Single memory bank shared across all benchmarks and tasks
- **Retrieval Pipeline**:
  - Embedding: `gemini-embedding-001`
  - FAISS IVF-Flat
  - Near-perfect retrieval accuracy
- **Self-evolution**:
  - After each task, the agent self-evaluates and writes:
    - Successful strategies worth reusing,
    - Recurring failure patterns to avoid.

New Memory Analysis Metrics

- **Retrieval Precision@1/@3:** Relevant retrieved memories
  - 98.6% P@1, 97.0% P@3
- **Utility Rate:** Memories influenced final reasoning.
  - 15% reveals critical retrieval-utilization gap
- **Quality Index:** success rate + efficiency + similarity
  - Mean = 0.392, range = [0.10, 0.58]
- **Quantity Ablation:** performance vs. memory bank size.
  - Best at ∼50 memories; decline beyond 150
- **Difficulty-Stratified Accuracy:**
  - Hard tasks: **+300%** (n=480)
  - Easy 25% to 37% (n=12)

Key Findings

- **Retrieval-Utilization Gap**: 98.6% precision vs 15% utilization
- **Task-Difficulty Specialization:** Hurts easy ones (overhead); helps where baseline struggles
- **Optimal Memory Bank Design:**
  - Using a single bank across tasks, demonstrates cross-task generalization and knowledge transfer
  - Quality matters: filter at >0.50
- **Efficiency Gains:**
  - Steps-to-success reduced by 41–46% when memory contributes

Conclusions & Next Steps

- Implementation of ReasoningBank and Memory-augmented agents:
  - Triple success rate on hard tasks (+300% on 97.6% of dataset)
  - Reduce interaction steps by 41–46%
- Analysis reveals:
  - Critical retrieval-utilization gap
  - Difficulty-dependent benefit
  - Clear optimal memory bank size

Future Directions:

- Adaptive retrieval based on task difficulty prediction
- Improved memory-conditioned action integration
- Memory consolidation and pruning