中國人民大學
RENMIN UNIVERSITY OF CHINA

Generalizing Experience for Language Agents with Hierarchical MetaFlows

Shengda Fan

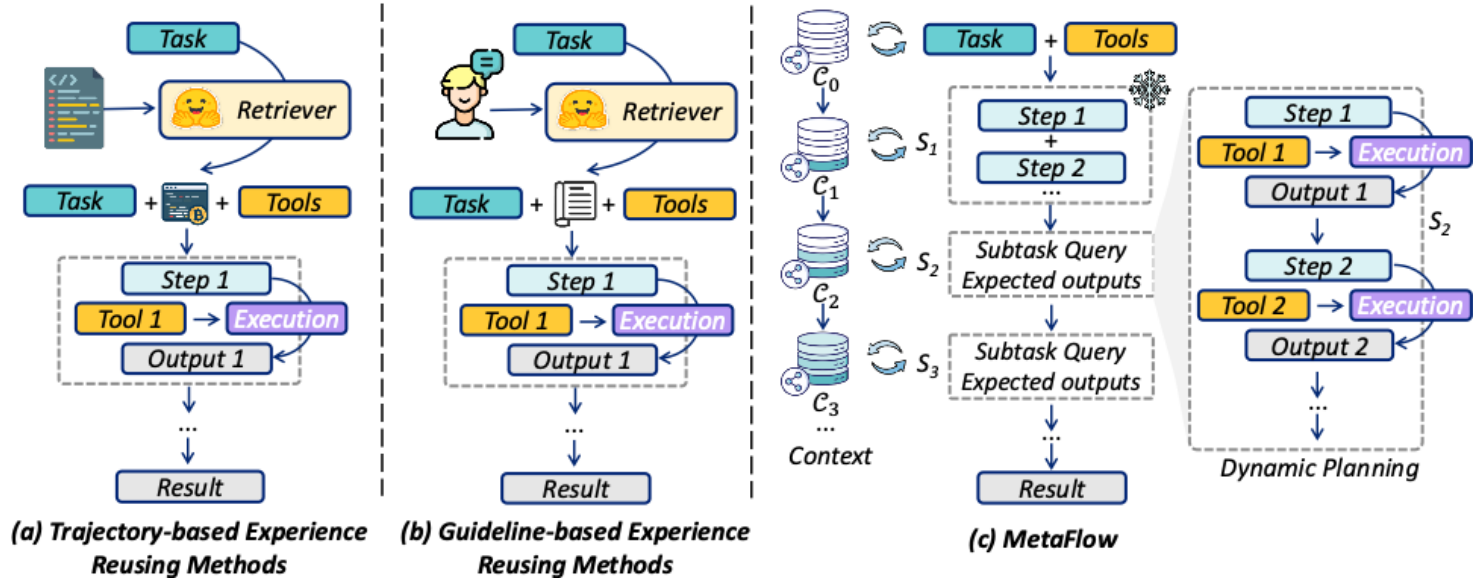
2025.11.07

Introduction

Problem: Existing agent experience reuse methods insert historical trajectories or natural language guidance into LLM prompts, which may introduce irrelevant details or miss key procedural steps.

Goal: Introduce a tree structure to organize historical experience and refine a generalizable MetaFlow through layer-by-layer abstraction for efficient task matching and reuse.

Core Idea: Define an executable MetaFlow structure, train a MetaFlowGen generator with verifiable reward signals, and build an experience tree using a hierarchical merging algorithm for efficient task reuse.

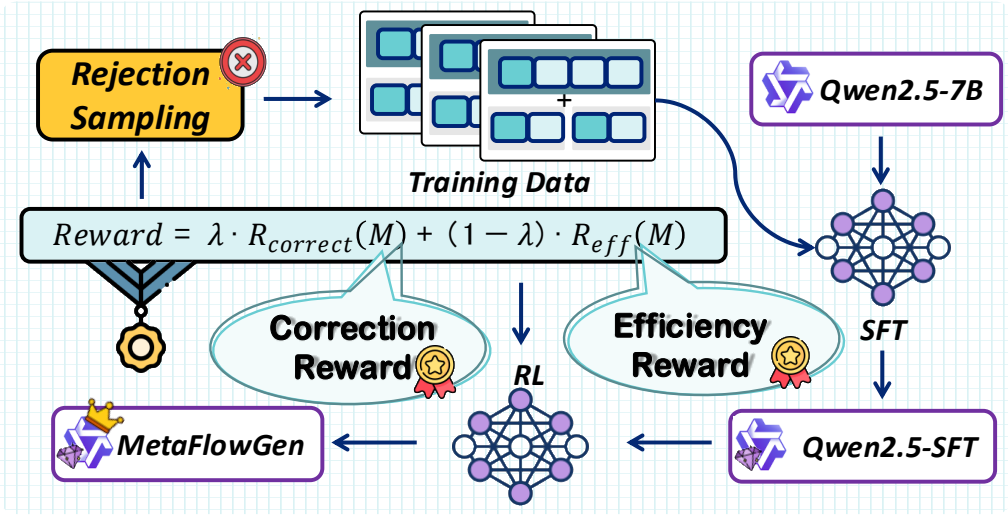


Methodology

- **The input to MetaFlowGen consists of task descriptions, MetaFlow, and tool descriptions for two samples, while the output is the MetaFlow and meta-task.**

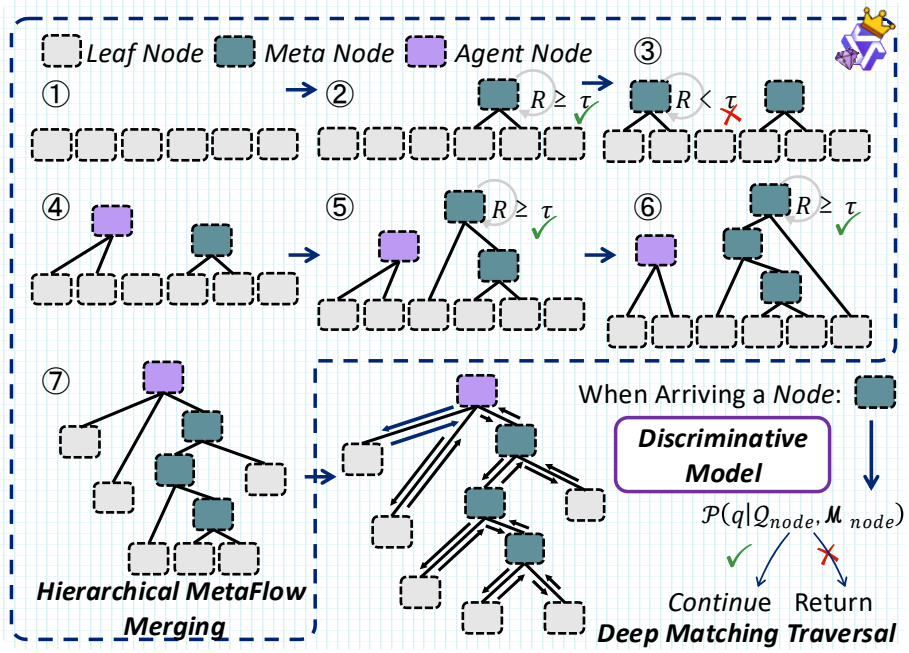
$$\mathcal{M}, \mathcal{Q} = \pi_\theta\left(q_1, q_2, m_1, m_2, \mathcal{A}\right)$$

- **This approach uses reinforcement learning to train the MetaFlowGen.**
  The training process is divided into two stages:
- **SFT cold-start using data obtained from rejection sampling.**
- **GRPO directly optimizes the model's reward.**
  The reward for each sample is the weighted sum of correctness (correct) and efficiency (Efficiency) rewards.



Methodology

- **To improve inference efficiency, this approach first inductively organizes historical data into an experience tree. Inspired by hierarchical clustering, we iteratively merge the two most similar nodes and remove low-quality nodes until a complete tree structure is formed.**
- **During inference, starting from the root node, the deepest node capable of solving the task is identified, and the corresponding MetaFlow is instantiated.**



The Process of MetaFlowLLM

Algorithm 1 Hierarchical MetaFlow Merging

Require: Leaf nodes $\mathcal{D} = \{(q_i, m_i)\}$, distance metric $d(\cdot, \cdot)$, MetaFlowGen θ, correctness threshold τ
Ensure: Hierarchical experience Tree $\mathcal{T}$

Initialize: Initialize the set of nodes $\mathcal{N} \leftarrow \mathcal{D}$
while $|\mathcal{N}| > 1$ **do**
  Select the most similar node pair: From $\mathcal{N}$, select two most similar nodes n_1, n_2 based on the distance metric $d(\cdot, \cdot)$
  Generate candidate MetaFlow: Merge n_1 and n_2 using the MetaFlowGen θ to produce a candidate MetaFlow $\mathcal{M}$
  Evaluate its correctness: Compute the correctness reward R_{correct} for $\mathcal{M}$
  if $R_{\text{correct}} >= \tau$ **then**
    Use $\mathcal{M}$ as the new parent node to merge n_1 and n_2
    Remove n_1, n_2 from $\mathcal{N}$ and add $\mathcal{M}$ to $\mathcal{N}$
  else
    Use pure-agent node as the new parent node to merge n_1 and n_2
    Remove n_1, n_2 from $\mathcal{N}$
  end if
end while
Final merge: If there are more than one pure-agent nodes, merge them into a single top-level pure-agent node
return The complete experience Tree $\mathcal{T}$

Experiments

- **On AppWorld, MetaFlow's TGC improvement is 33.1% and 31.6%, outperforming all baseline methods.**
- **On WorkBench, MetaFlow improves accuracy by an average of 6.16% and reduces side effects by 4.13%, outperforming all baseline methods.**

Model	Agent Type	Configuration			
		Base	*w/ Traj.*	*w/ Guideline*	*w/ MetaFLow*
Qwen2.5-7B	Reflexion	7.6	32.7	13.5	**37.4**
	ReAct	6.4	37.4	17.0	**43.9**
Qwen2.5-32B	Reflexion	12.9	33.9	22.8	**43.9**
	ReAct	22.2	49.1	38.6	**50.9**
GPT-4o-mini	Reflexion	13.5	37.4	20.5	**45.6**
	ReAct	7.0	22.2	7.0	**35.7**

Model	Analytics		Calendar		CRM		Email		PM		MD		Avg	
	%acc↑	%se↓	%acc↑	%se↓	%acc↑	%se↓	%acc↑	%se↓	%acc↑	%se↓	%acc↑	%se↓	%acc↑	%se↓
GPT-4o	30.8	53.8	53.0	19.7	30.9	14.5	32.8	25.9	8.16	0.0	16.3	48.9	27.2	33.1
Qwen2.5-7B	18.2	60.3	27.6	**35.5**	16.0	**30.5**	9.31	**43.8**	8.57	**6.12**	9.39	**53.9**	14.3	**43.0**
w/ Traj.	**30.8**	46.2	18.2	48.5	18.2	32.7	**15.5**	46.6	**14.3**	12.2	8.16	57.1	16.3	44.8
w/ Guideline	21.8	51.3	**37.9**	47.0	20.0	52.7	15.5	55.2	10.2	6.12	8.16	60.5	**17.4**	49.5
w/ MetaFlow	26.9	**43.6**	25.8	50.0	**23.6**	34.5	10.3	55.2	14.3	18.4	**9.52**	57.8	17.2	46.8
Qwen2.5-32B	6.41	88.5	57.6	24.2	12.7	43.6	27.6	**24.1**	6.12	**0.0**	12.2	61.9	19.2	47.2
w/ Traj.	**25.6**	70.5	37.9	40.9	**41.8**	**20.0**	8.62	58.6	**20.4**	12.2	15.0	60.5	23.2	49.0
w/ Guideline	19.2	79.5	59.1	21.2	27.3	23.6	31.0	36.2	10.2	2.0	15.0	57.8	25.2	43.3
w/ MetaFlow	24.4	**69.2**	**59.1**	**19.7**	27.3	23.6	**37.9**	32.8	12.2	4.1	**18.4**	**55.8**	**28.3**	**40.4**
GPT-4o-mini	11.5	44.9	48.5	30.3	21.8	36.4	24.1	58.6	8.16	6.12	12.2	56.5	19.6	43.0
w/ Traj.	12.8	44.9	39.4	25.8	27.3	36.4	20.7	48.3	12.2	12.2	14.3	57.1	19.9	41.9
w/ Guideline	12.8	43.6	47.0	37.9	**27.3**	30.9	31.0	51.7	12.2	4.10	12.9	60.5	21.9	43.5
w/ MetaFlow	**19.2**	**30.8**	**53.0**	**15.2**	16.4	**23.6**	**31.0**	**46.6**	**16.3**	**4.08**	**22.4**	**51.7**	**26.1**	**33.6**

Thank You!