

Memento: Fine-tuning LLM Agents without Fine-tuning LLMs

Huichi Zhou^{*1,2}, Yihang Chen^{*2}, Siyuan Guo³, Xue Yan⁴, Kin Hei Lee, Zihan Wang, Ka Yiu Lee²,
Guchun Zhang², Kun Shao², Linyi Yang^{†2}, and Jun Wang^{†1}

¹AI Centre, UCL, ²Huawei Noah's Ark Lab, UK, ³Jilin University, ⁴Institute of Automation, CAS

代码: <https://github.com/Agent-on-the-Fly/Memento>

简介

本文提出Memento，一个基于planner-executor agent架构deep research方案，最大特点是集成了基于记忆的持续学习机制，使agent能够在不tuning llm参数的情况下实现self-evolution。不同于依赖固定工作流的静态agent或通过频繁tuning底层llm参数的高成本部署方案，Memento借鉴人类 episodic memory(情节记忆)与case-based reasoning(基于案例推理)的机制，将过往历史任务执行轨迹存储为外部记忆，并通过检索与复用相似案例来指导新任务的决策。最值得关注的是，Memento引入了soft Q-learning学习如何检索最有价值的记忆。

背景

本文属于self-evolution agents/memory/deep research方向的工作。目前有两种agents部署范式：一是静态agent，即根据固定的workflow来执行任务，虽然行为可控但缺乏灵活性和泛化性；二是通过tuning agent底层的llm参数来提升性能，具有一定的适应新环境的灵活性，但是tuning的成本比较高，特别是在动态多变的环境中部署后如果经常tuning成本可能吃不消。

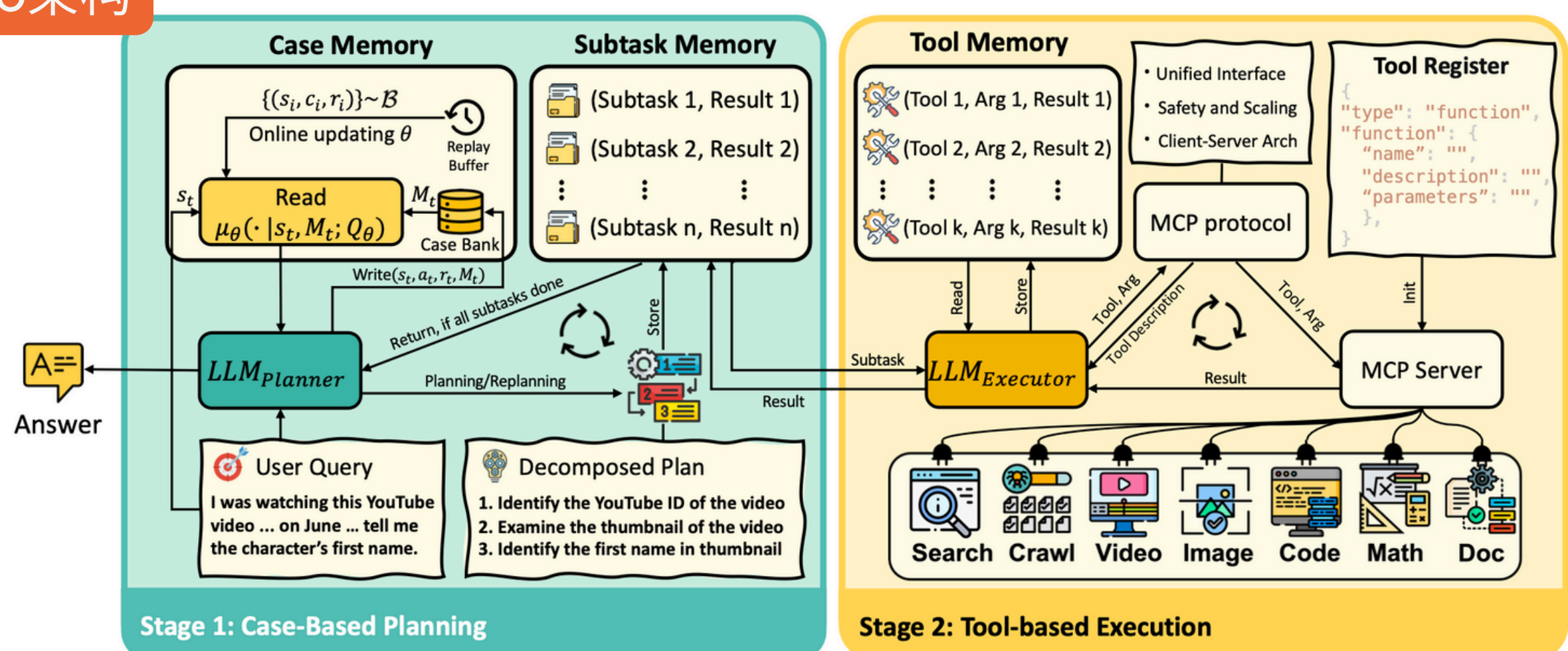
作者受到人类学习机制的启发：存储过往经历的episodic memory(情节记忆)，在遇到新问题通过 类比推理(case-based/analogy-based reasoning) 将旧经验迁移到新任务中，从而快速解决新任务。本文提出了一种基于记忆(memory)驱动的学习框架，旨在让agents在不改变底层llm参数的前提下，通过不断积累和复用任务轨迹，实现低成本的continual learning(持续学习)与self-evolution。

实验设置

具体实现方面，论文提到Memento集成了三个memory模块

- Case Memory (vectorised storage of prior cases for high-level planning)
- Subtask Memory (text-based storage of active subtasks and their results)
- Tool Memory (text-based logs of tool interactions for each subtask)

Memento架构



思考

- 1.这篇论文昨天刚挂到Arxiv的时候叫AgentFly，今天改名Memento就直观多了
- 2.简单理解，就是实现了一种memory机制，我们之前也读过mem0、mem1、memoryos等工作，本文最大特点是不再把memory机制看作纯工程设计，而是引入了soft q-learning进行检索优化。当然至于这类学术创新是否有效果，有待检验吧
- 3.目前开源代码还没包含CBR相关内容，所以对于论文我也只是简单看了看，毕竟读不懂的地方也没代码参考