面向开放环境的大模型持续学习研究

周 杰 青年研究员
计算机科学与技术学院

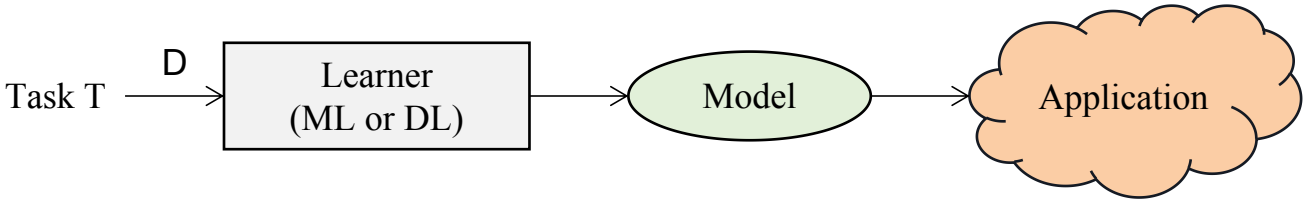
華東師範大學
EAST CHINA NORMAL UNIVERSITY

華東師範大學
EAST CHINA NORMAL UNIVERSITY

目录 | CONTENT

持续学习背景

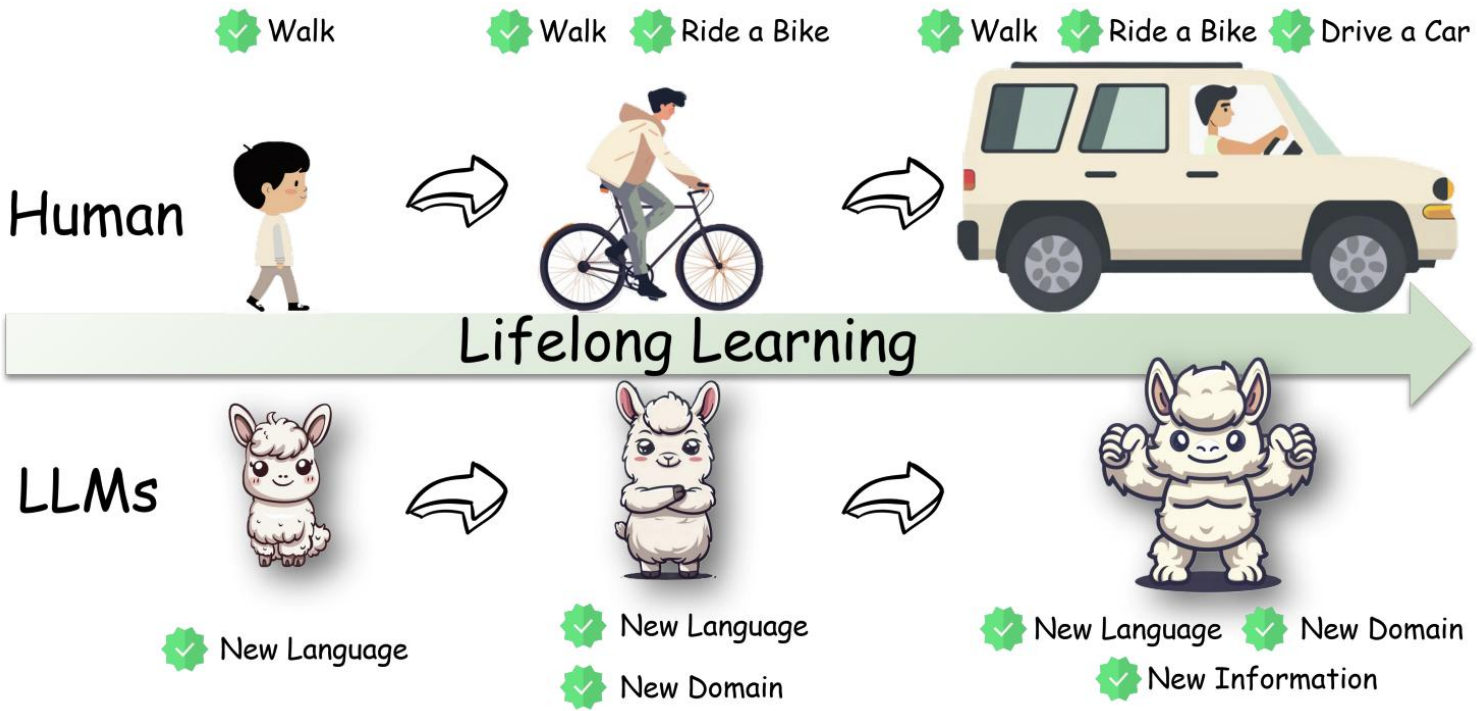
- 传统机器学习模型: 一个任务一个模型



- 存在主要不足
  - **封闭环境假设**：不会出现新的东西
    - 开放世界假设：会出现不知道的或者新的东西，OOD问题
  - **没有持续学习**：不会有知识获取或者迁移
  - **部署后模型不再学习**：模型部署后模型参数固定

Liu, Bing. "Learning on the job: Online lifelong and continual learning." AAAI. 2020.

持续学习背景

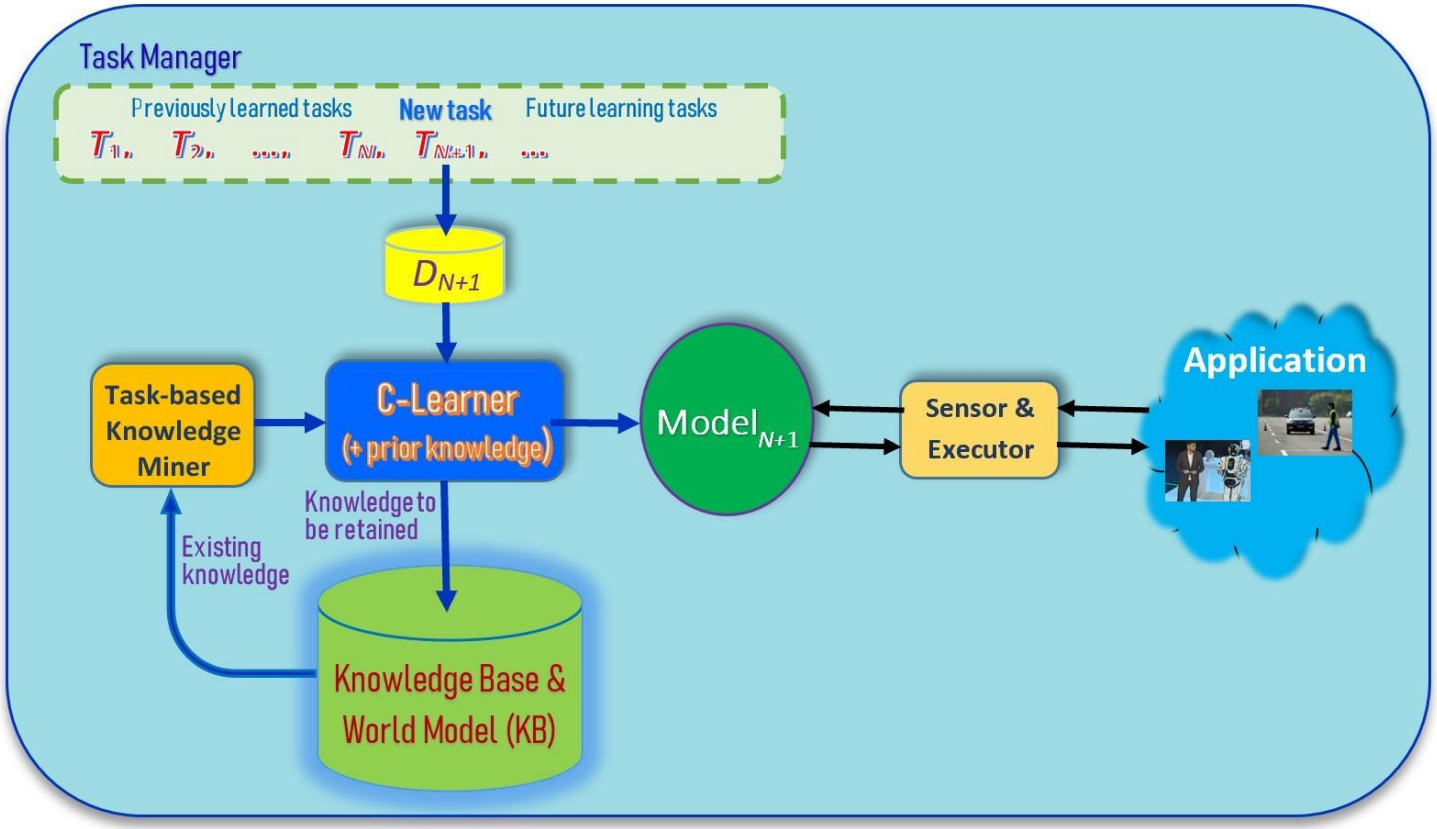


Junhao Zheng, et al. Towards Lifelong Learning of Large Language Models: A Survey

传统持续学习

- 封闭环境持续学习模型: 一个模型学习一系列任务
  - Continual learning/Lifelong Learning/Increment Learning
- 任务定义：依次学一连串任务$T_1, T_2, \ldots, T_N, \ldots$. 每一个任务t都包含一个完整训练数据.
  - 克服灾难性遗忘（catastrophic forgetting，CF）： 学习新的任务T_{N+1}而不遗忘以前N个任务的能力
  - 知识迁移（knowledge transfer，KT）：利用前面任务学习的知识用于学习新的任务T_{N+1}
    - 正向迁移
    - 反向迁移

传统持续学习



假设:

当一个任务学完以后，该任务的数据不可获取，至少大部分可以获取

任务T_{N+1} 和数据集 D_{N+1} 都是完整给定的

$Y_{test} \in Y_{train}$

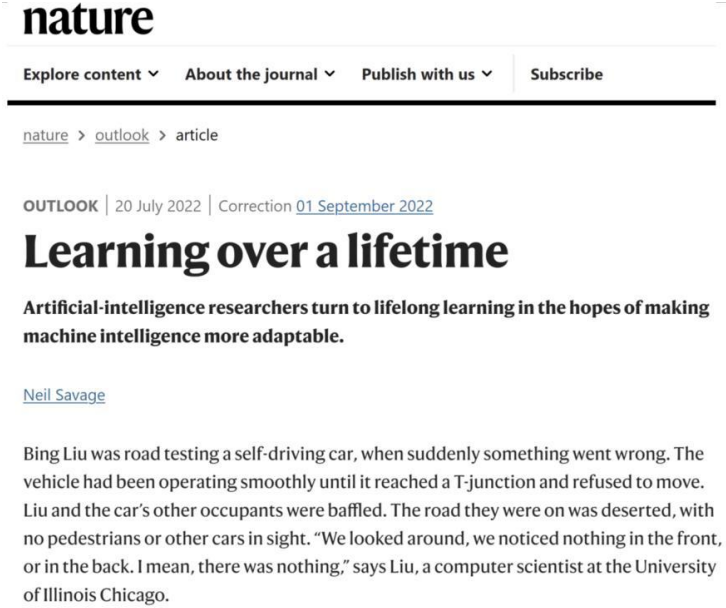
Liu. Learning on the Job: Online Lifelong and Continual Learning. AAAI-2020

开放环境持续学习

- 开放环境持续学习
  - (1) 识别新的物体 (OOD 识别)
  - (2) 在**标注以后**增量学习新的物体

以自动驾驶公司为例

- 一辆车在路上做测试
  - 在一个十字交叉口，车停止了并拒绝前进
- 人类驾驶员接手
  - 传感器发现路上有一个鹅软石，车说 ”我发现一个不知道的东西，我该怎么办？”
  - 应该回复 “是安全的，继续前进”
- 车辆学习到这个物品同时下次不会在有问题



nature

Explore content ∨ About the journal ∨ Publish with us ∨ Subscribe

nature › outlook › article

OUTLOOK | 20 July 2022 | Correction 01 September 2022

Learning over a lifetime

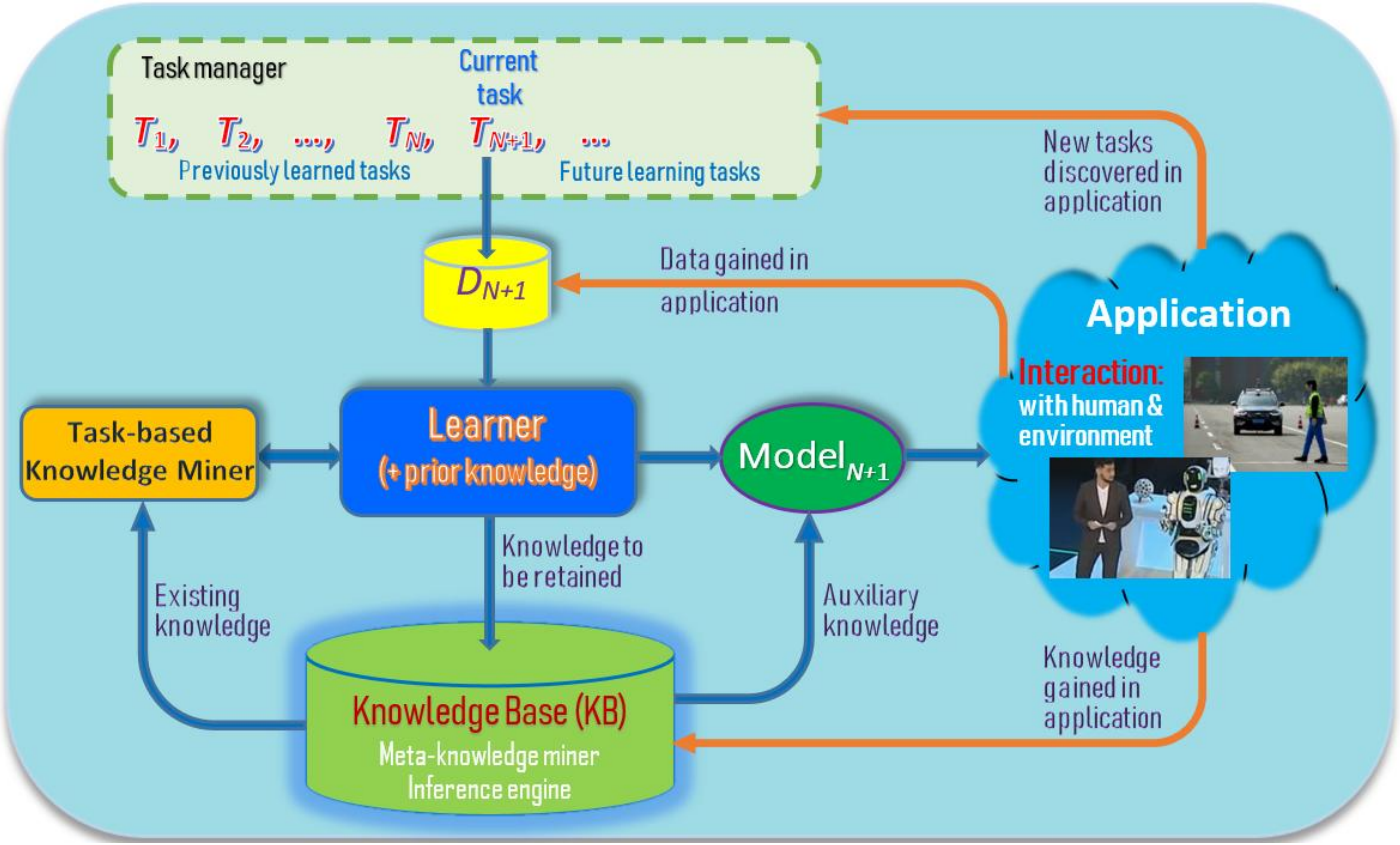
Artificial-intelligence researchers turn to lifelong learning in the hopes of making machine intelligence more adaptable.

Neil Savage

Bing Liu was road testing a self-driving car, when suddenly something went wrong. The vehicle had been operating smoothly until it reached a T-junction and refused to move. Liu and the car's other occupants were baffled. The road they were on was deserted, with no pedestrians or other cars in sight. "We looked around, we noticed nothing in the front, or in the back. I mean, there was nothing," says Liu, a computer scientist at the University of Illinois Chicago.

Liu, Bing. "Learning on the job: Online lifelong and continual learning." AAAI. 2020.

开放世界持续学习



橘色的线:

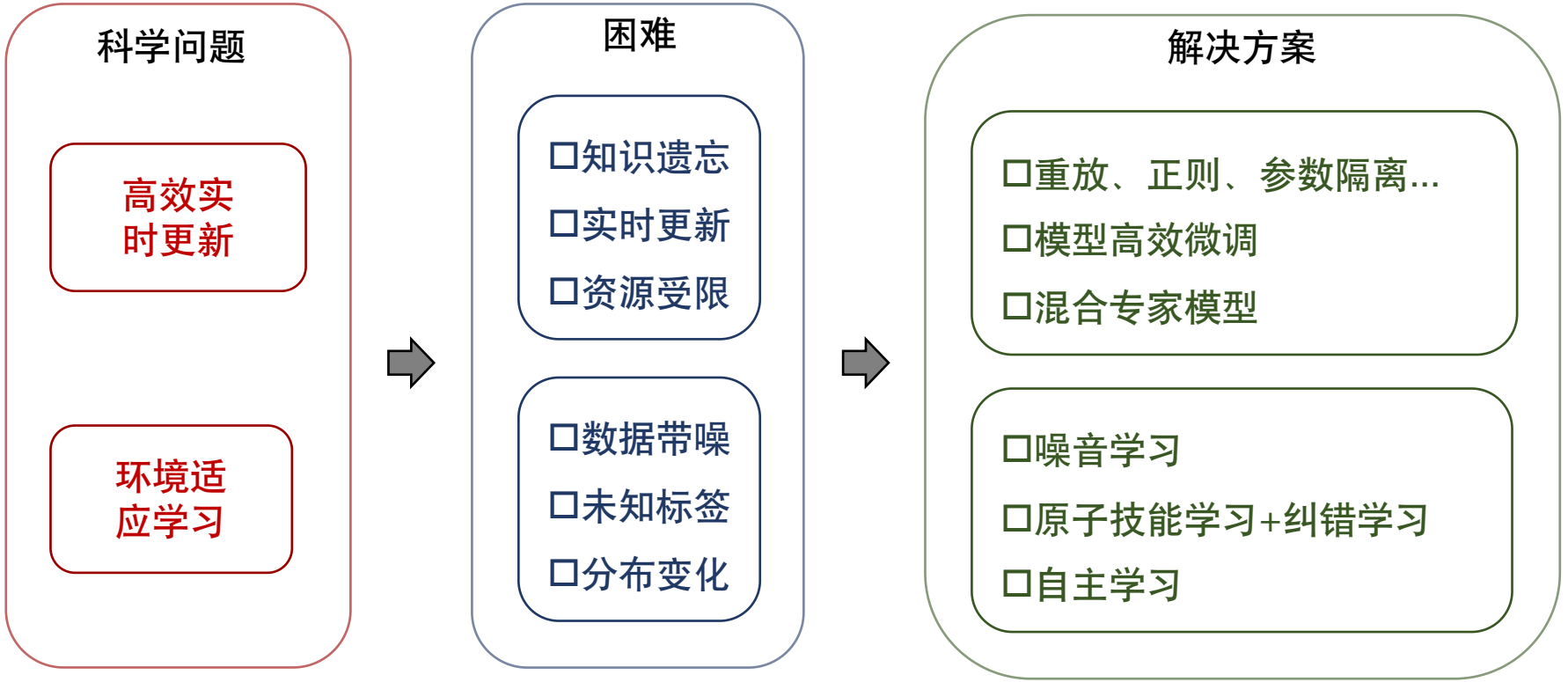
模型部署后学习-在工作中学习

Liu, Bing. "Learning on the job: Online lifelong and continual learning." AAAI. 2020.

華東師範大學
EAST CHINA NORMAL UNIVERSITY

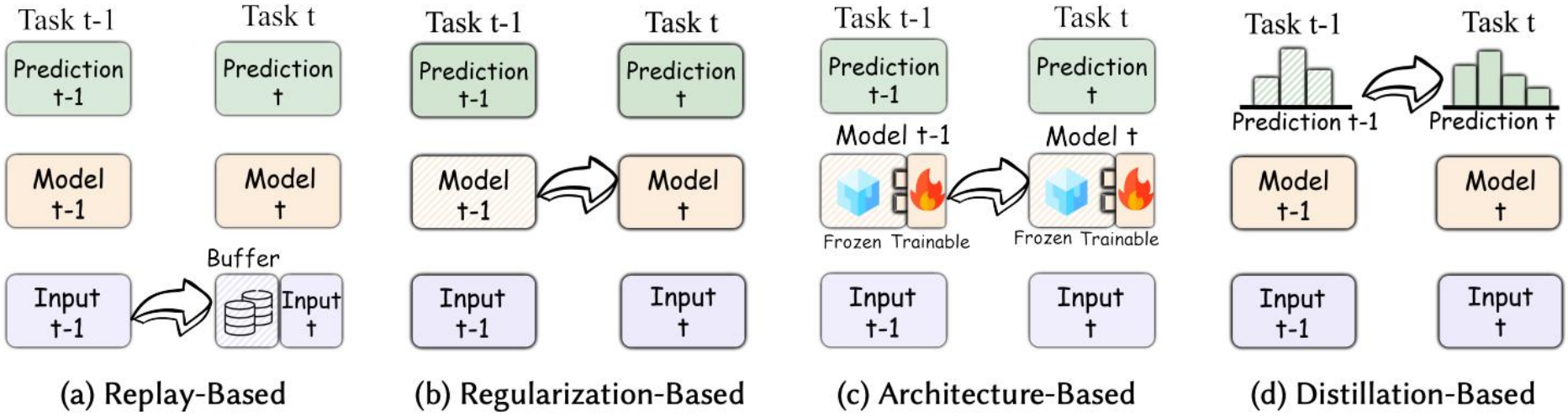
目录 | CONTENT

开放环境大模型持续学习挑战



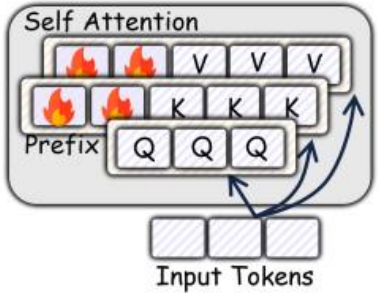
持续学习中遗忘问题

- 常见的四种持续学习方案
  - **数据回放**：采样少量历史任务数据进行训练
  - **基于正则**：利用正则对参数进行约束，不偏离太远
  - **基于架构**（如：参数隔离）：不同任务之间参数独立
  - **基于蒸馏**：历史模型作为一部分引导



基于高效微调的持续学习

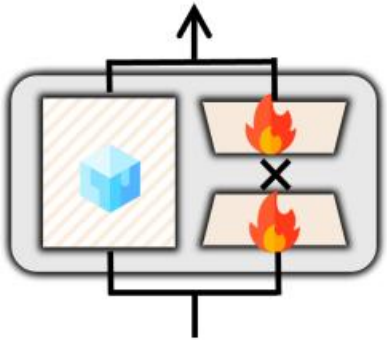
(a) Prompt Tuning



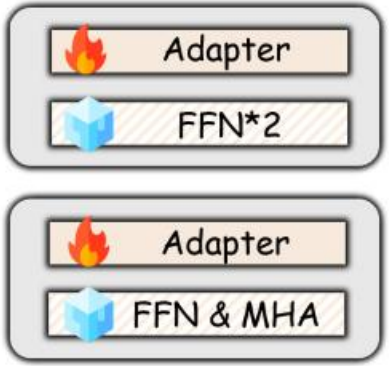
(b) Prefix Tuning

传统模型主要采用额外添加参数的方式

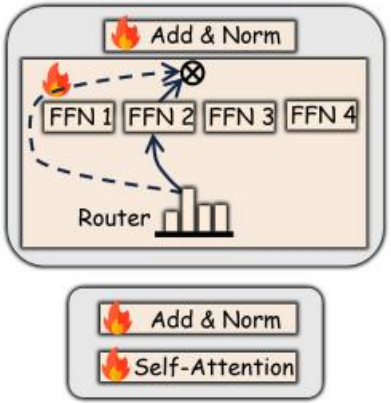
如何精准定位参数进行修改?



(c) LoRA



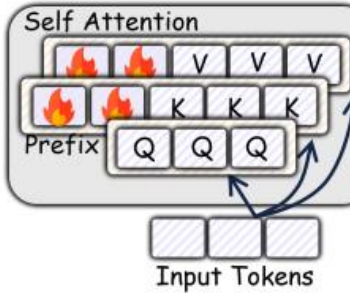
(d) Adapters



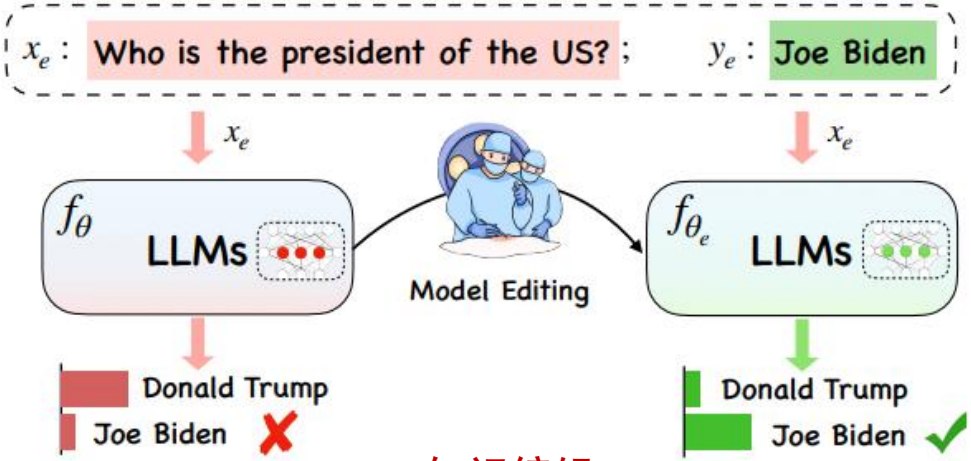
(e) Mixture of Experts

基于高效微调的持续学习

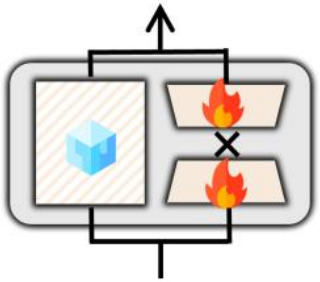
(a) Prompt Tuning



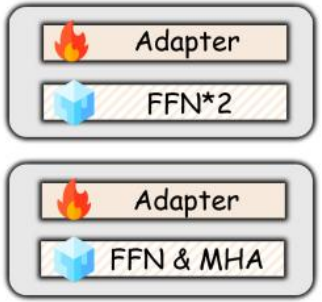
(b) Prefix Tuning



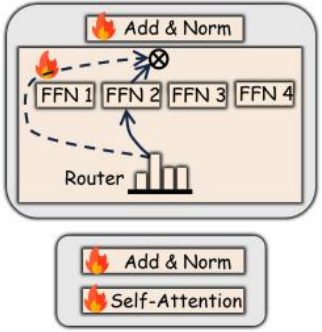
知识编辑



(c) LoRA



(d) Adapters



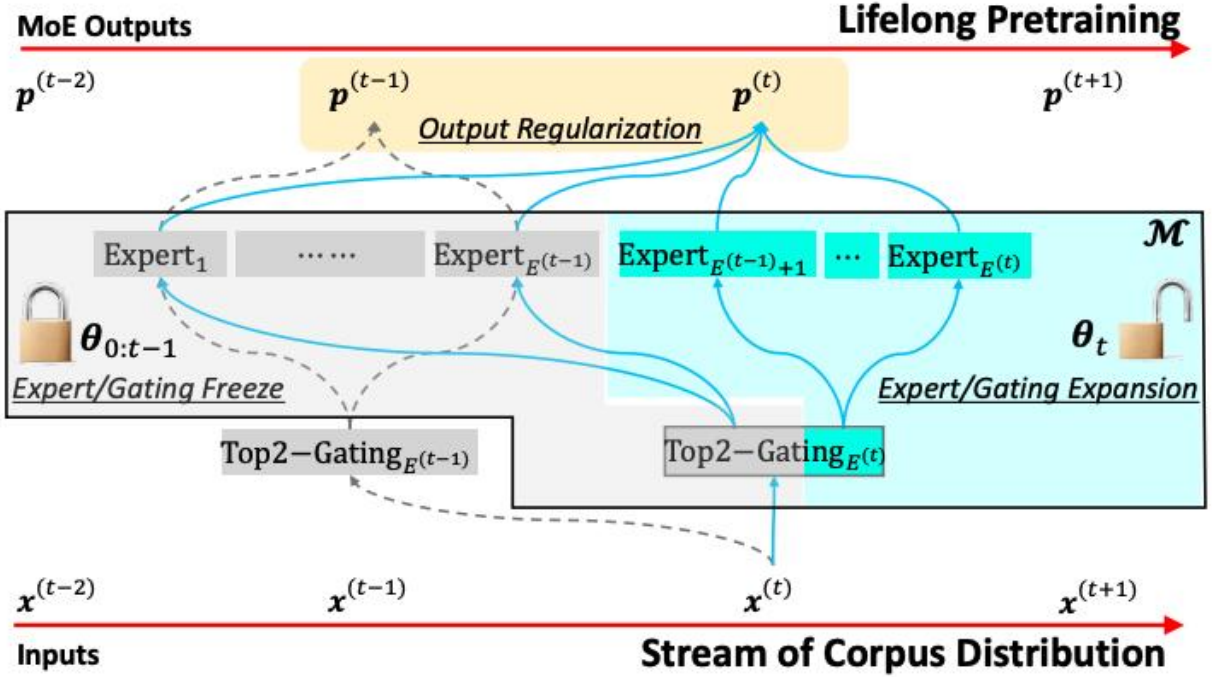
(e) Mixture of Experts

传统模型主要采用额外添加参数的方式

如何精准定位参数进行修改?

基于混合专家模型（MoE）的持续学习

- MoE用于大模型持续预训练
  - 不同专家代表不同能力

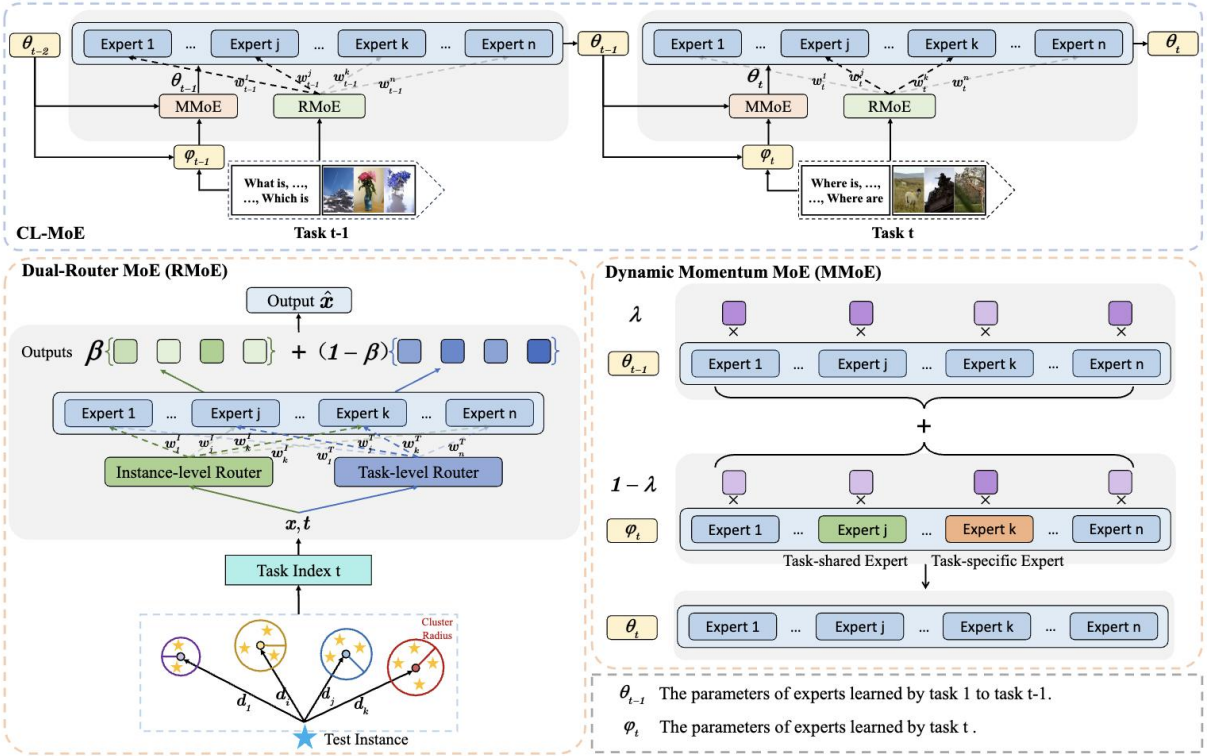


- 专家扩展：
  - 随着语料增加，专家和Gating也会随之扩展
  - 原来的专家和Gating不更新
- 隐式正则
  - 利用蒸馏的方式来进行正则优化，防止遗忘问题

Wuyang Chen, et al. Lifelong Language Pretraining with Distribution-Specialized Experts, ICML, 2023.

基于混合专家模型（MoE）的持续学习

- 不同任务共享专家、一个任务需要不同专家

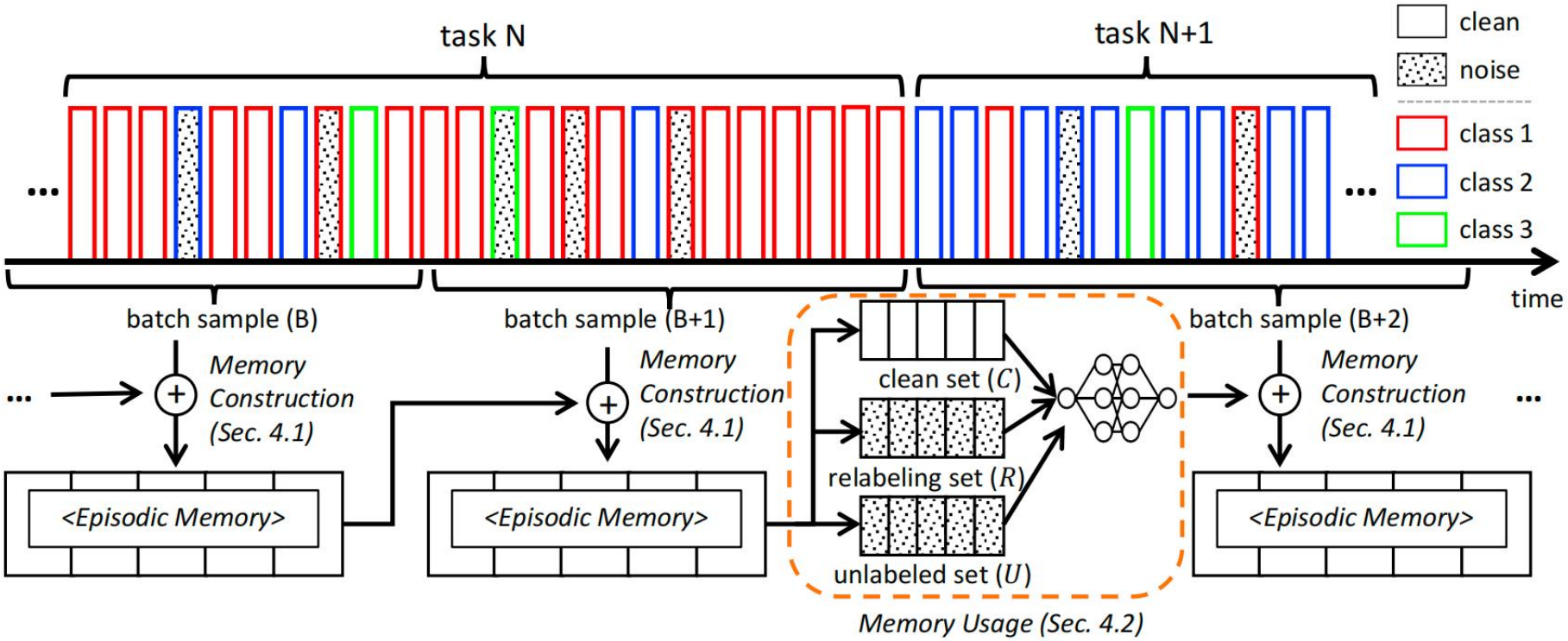


- 双Router机制：
  - 一个关注样本级别局部信息
  - 一个关注任务级别全局信息
- 动态动量更新
  - 对于任务共享专家和任务特定专家采用不同的更新方式

Tianyu Huai, et al. CL-MoE: Enhancing Multimodal Large Language Model with Dual Momentum Mixture-of-Experts for Continual Visual Question Answering, 2024.

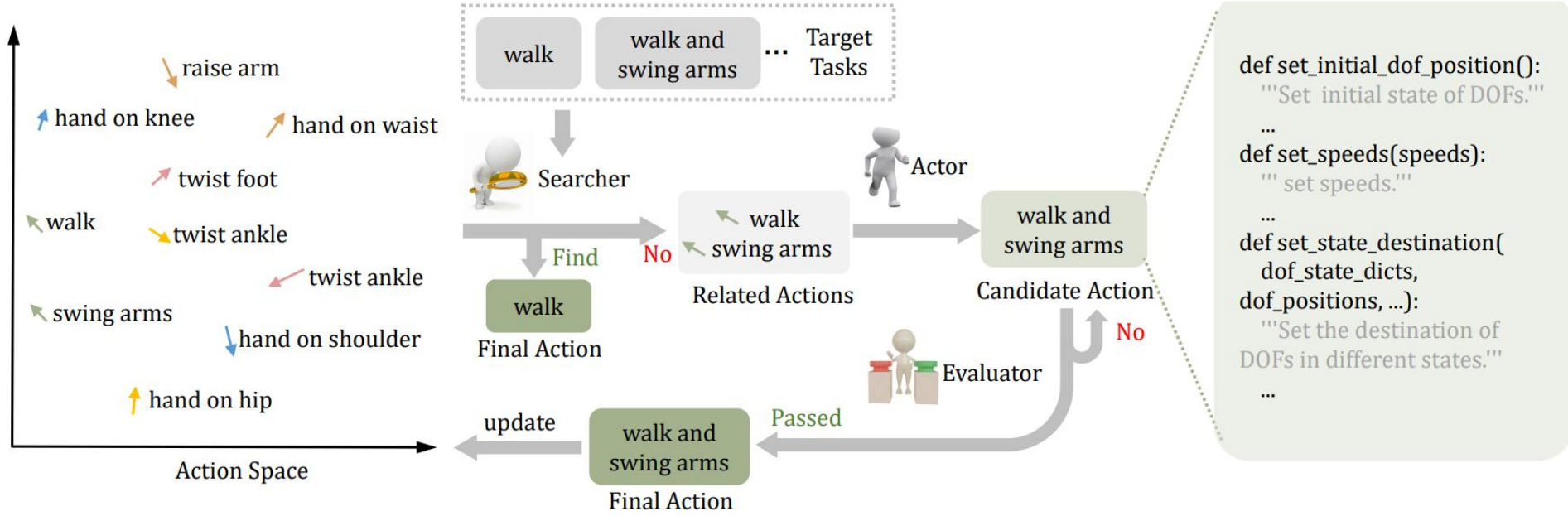
噪音数据下持续学习

- 传统：单个任务完整数据，且数据都是干净的
- 论文：数据流中数据标注是带噪音的



Chris Dongjoo Kim, et al. Continual Learning on Noisy Data Streams via Self-Purified Replay, ICCV, 2021
Jihwan Bang, et al. Online Continual Learning on a Contaminated Data Stream with Blurry Task Boundaries, CVPR, 2022

持续原子技能（Low-level）学习

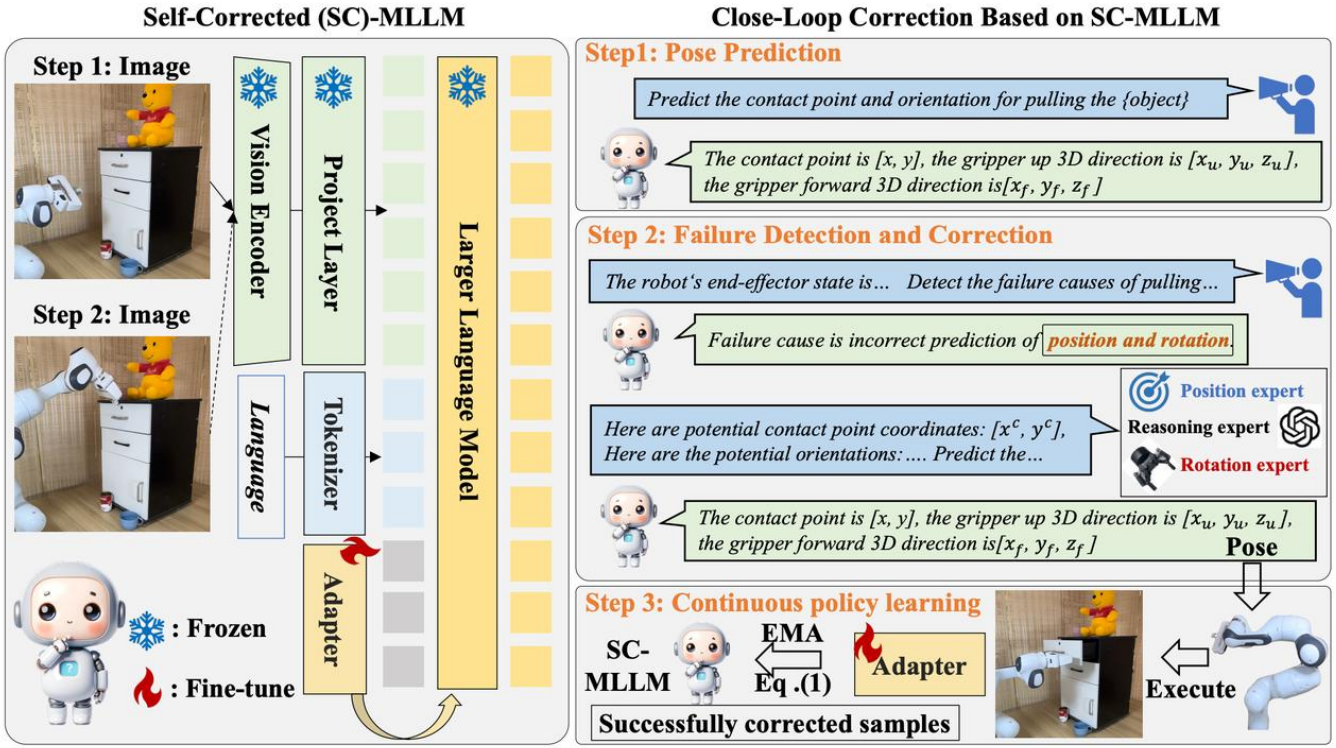


- **Searcher**：
  - 根据目标任务查询Action，从Action库里面搜索，大于一定阈值则选择该Action
  - 找不到，则新建一个Action，生成Action的动作
- **Evaluator**
  - 评测当前Action是否可行
  - 更新到知识库

RoboCoder: Robotic Learning from Basic Skills to General Tasks with Large Language Models, 2024

从纠正反馈中持续学习

- 面对新的任务或者实例对象时，缺乏从纠正反馈中学习的能力



犯一次错是无知，
犯两次错是愚蠢。

如下的步骤不断迭代：

1）多模态大模型生成操作，包含位置和方向

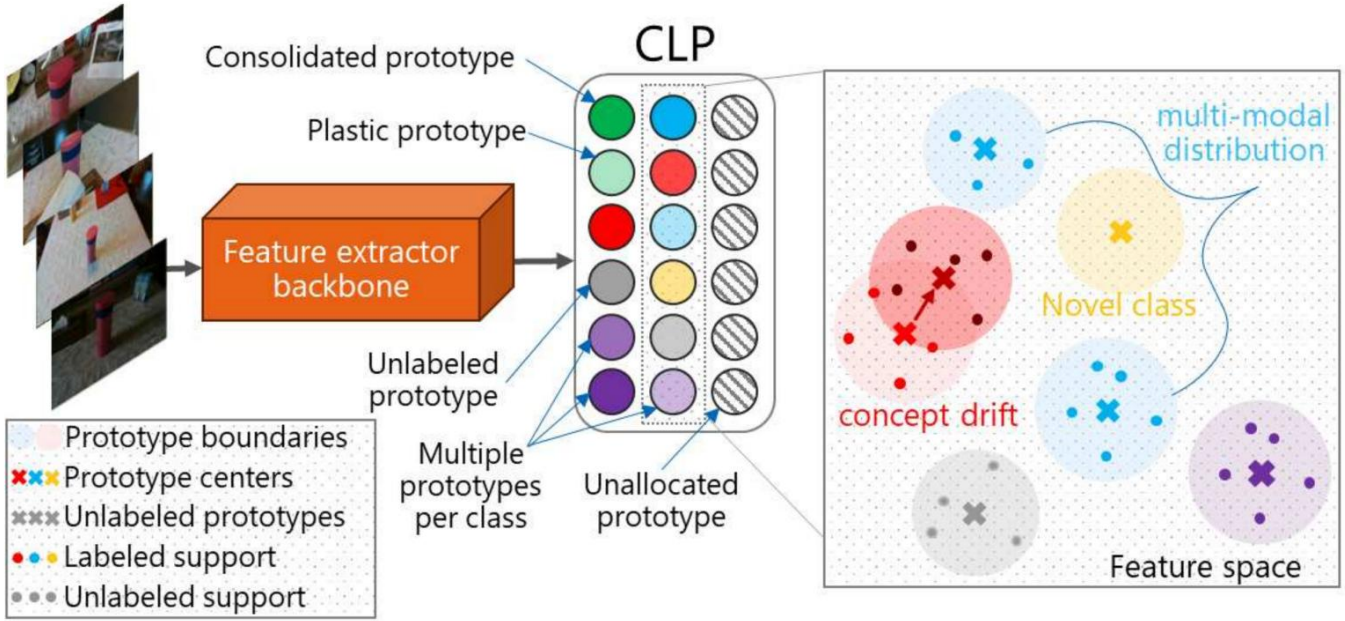
2）专家纠正模型预测结果

3）纠正后数据用于微调多模态大模型

Self-Corrected Multimodal Large Language Model for End-to-End Robot Manipulation

大模型自主持续学习

- 开放环境下的自主学习机器人：
  - 人类学习不仅包括少量直接指导，还包括大量未标记的经验，这种无监督学习是持续、自主、互动的
  - 对于t时刻，已知类别集合 1,2, ..., C，会出现没有标注过且属于未知类别C+1...的样本



学习类别原型表示

一个复杂类别可能包含多个原型表示

无标签数据进行聚类

当遇到新标签数据对聚类中心打上标签

Continual Learning for Autonomous Robots: A Prototype-based Approach, Arxiv, 2024.

華東師範大學
EAST CHINA NORMAL UNIVERSITY

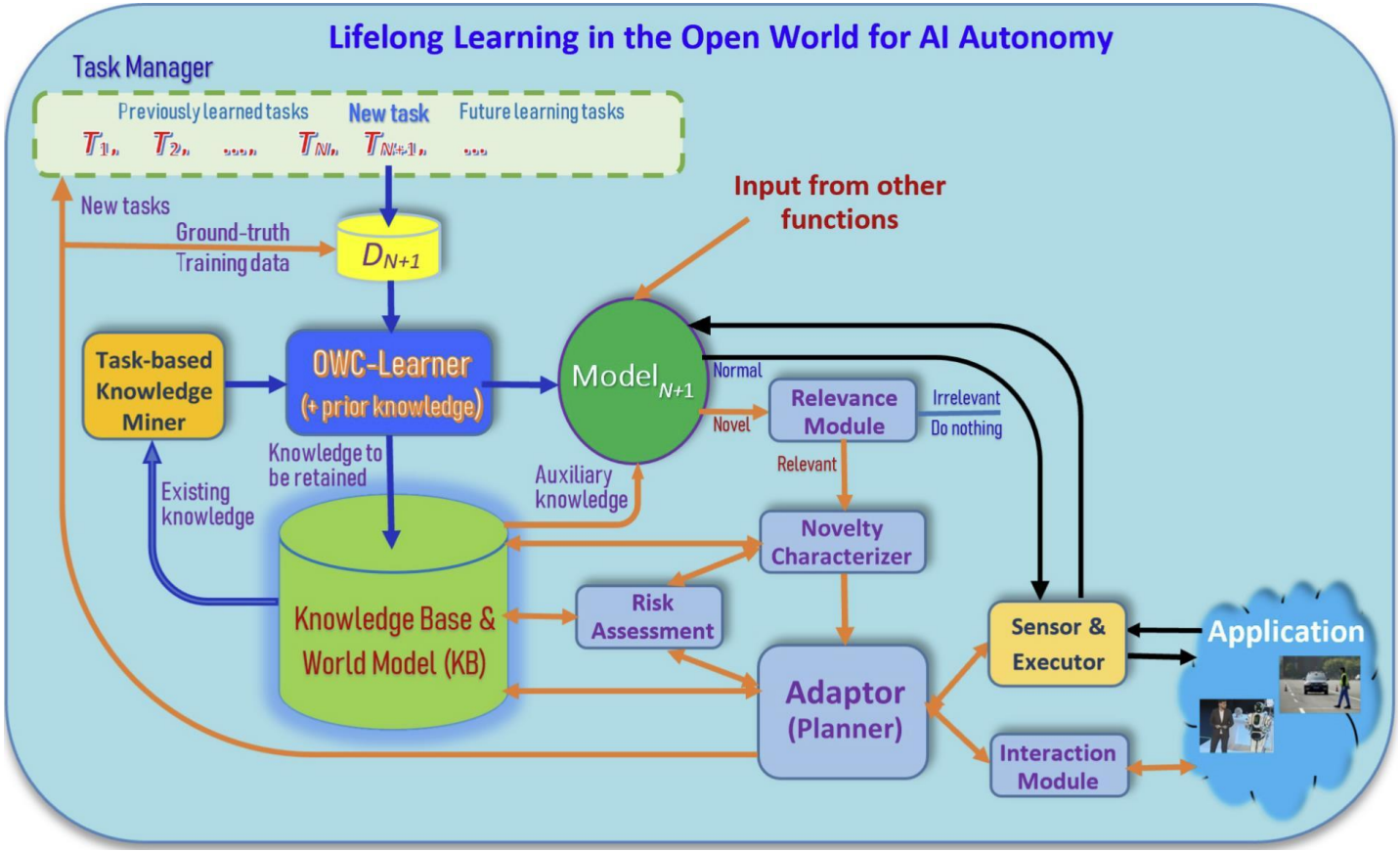
目录 | CONTENT

自主（Autonomous）持续学习

- 现实世界是开放和动态的环境，充满了未知，最终AI智能体需要自主学习开放环境持续学习（SOLA）
  - SOLA: 在模型部署后，模型自动和持续地使用自主和自我监督方式来适应为止的环境
- **自我驱动**: 发现并学习未知的物品
  - 好奇心是人类学习的内在驱力
- **自我监督**: 利用智能体自己收集训练数据
  - 和人、其他智能体和环境进行交互
- **适应**: 适应新的/分布外的环境
  - 需要计划、动作和风险评估

Liu, Mazumder, Robertson and Grigsby. AI Autonomy: Self-Initiated Open-World Continual Learning and Adaptation. AI Magazine, May 21, 2023

自主（Autonomous ）持续学习



Relevance: 是否和当前任务相关

Novelty Characterizer：刻画新的物品

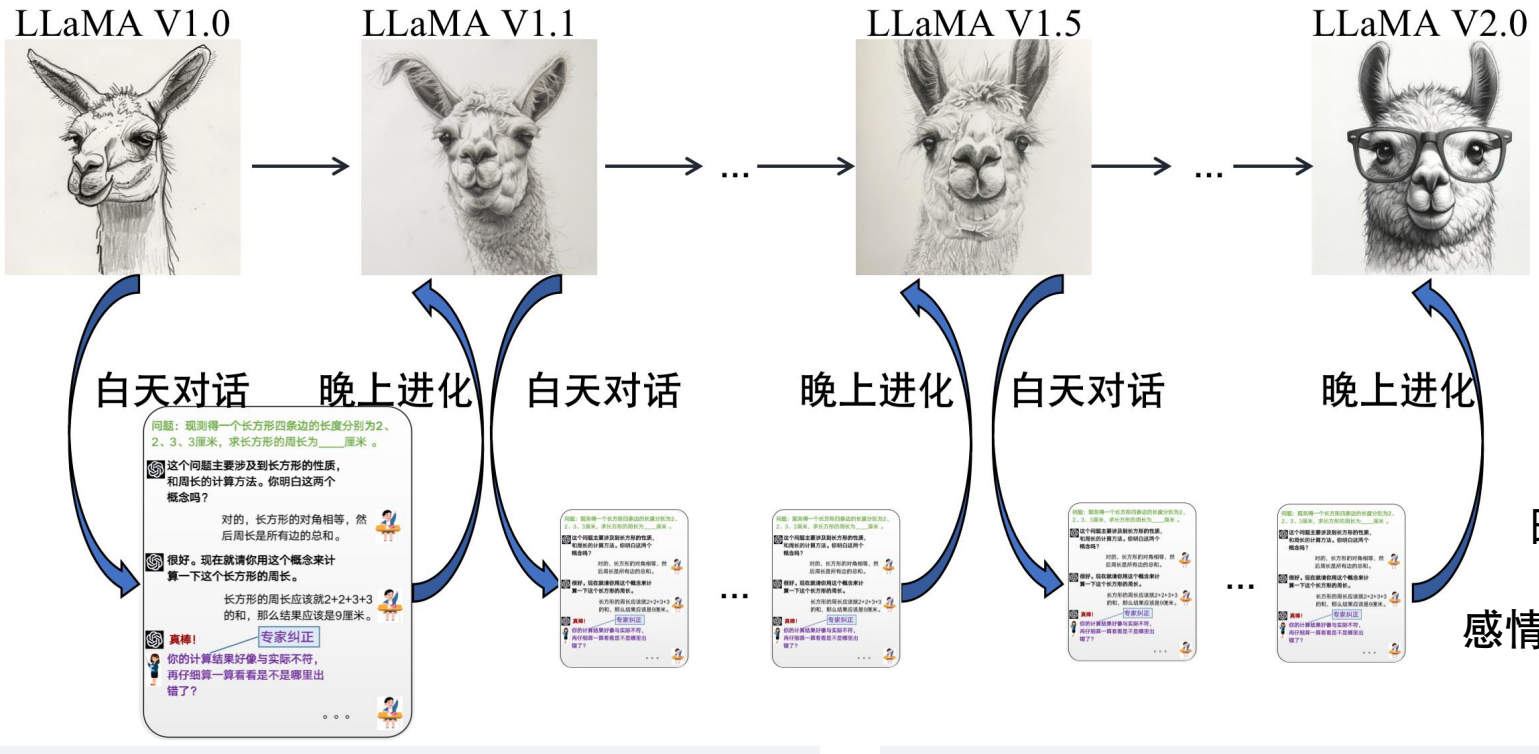
Adaptor：适应新的物品，并给出plan

Risk：刻画风险

Interaction: 交互，比如用来获得标注

Liu, Mazumder, Robertson and Grigsby. AI Autonomy: Self-Initiated Open-World Continual Learning and Adaptation. AI Magazine, May 21, 2023

可持续学习的个性化大模型



陪了三年的朋友还是三年前的那个朋友吗？

日久见人心

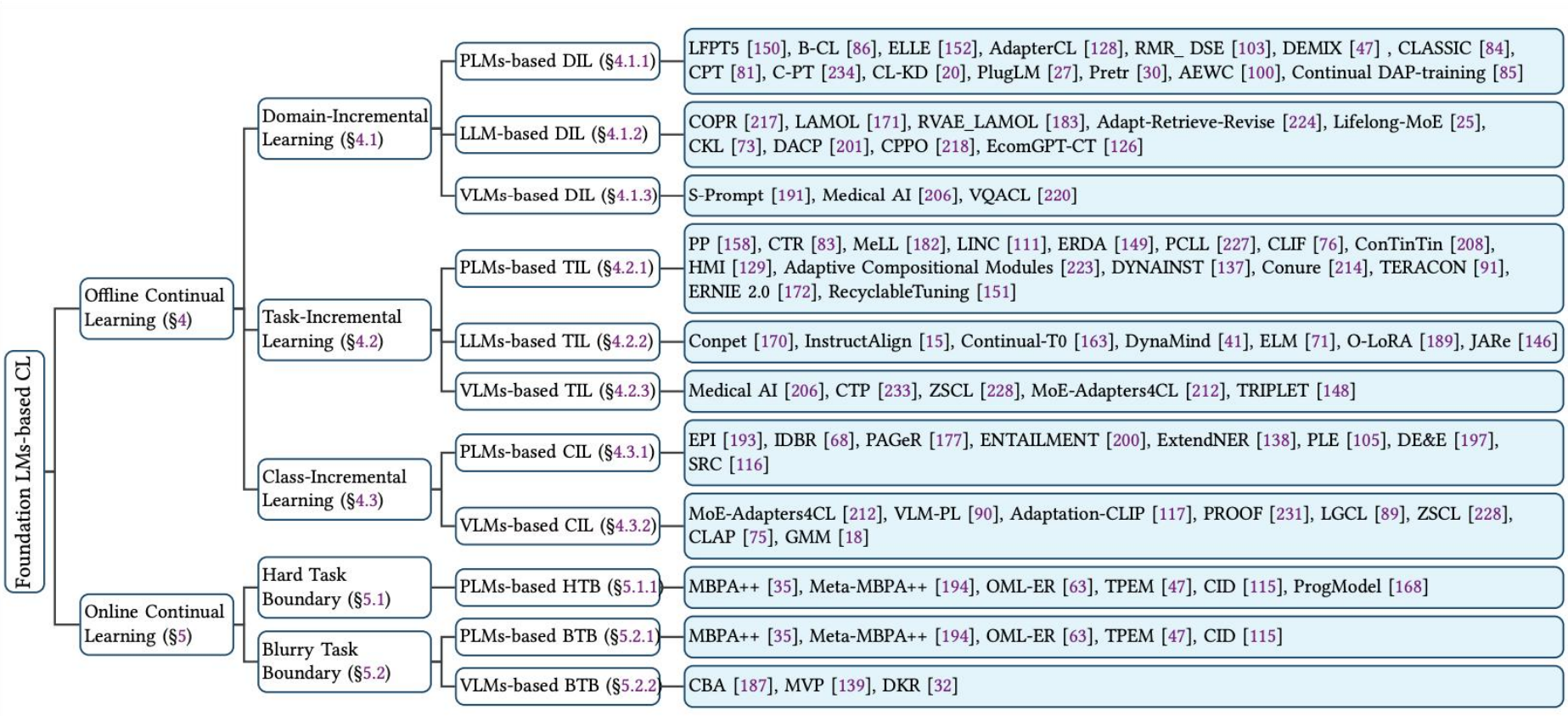
感情是可以培养的

大模型存在最大挑战：

- 缺少数据，尤其是过程交互数据
- 共用一个大模型，无法实现真正个性化

- 一边收集数据一边更新模型
- “长久陪伴”实现“一人一模型”
- 人机共同成长，实现“人机共生”

基于持续学习的情感分析



Yang, Yutao et al. "Recent Advances of Foundation Language Models-based Continual Learning: A Survey.", Arxiv, 2024.

谢谢！