



# 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

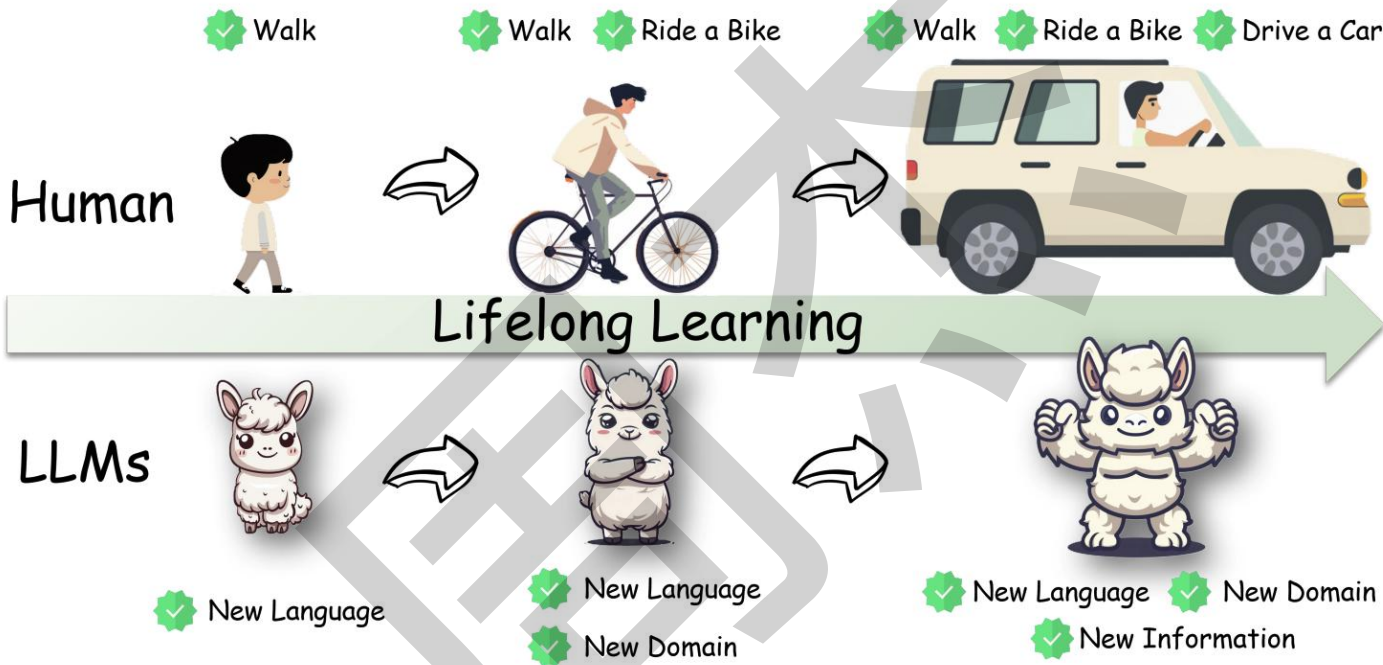


華東師範大學  
EAST CHINA NORMAL  
UNIVERSITY

## 目录 | CONTENT

- 持续学习背景介绍
- 开放环境大模型持续学习
- 持续学习未来趋势

# 持续学习背景

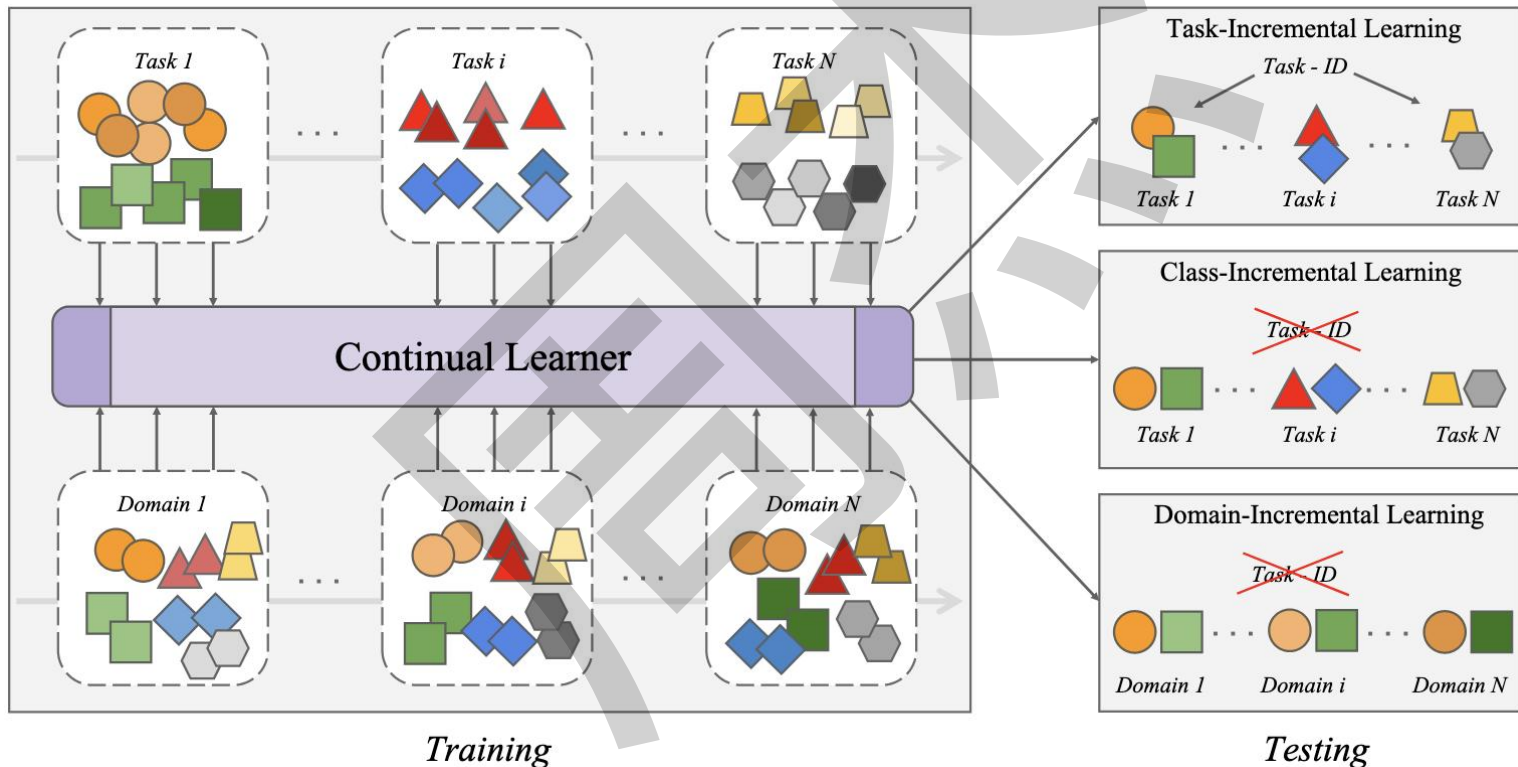


- 克服灾难性遗忘：学习新的任务 $T_{N+1}$ 而不遗忘以前 $N$ 个任务的能力
- 知识迁移：利用前面任务学习的知识用于学习新的任务 $T_{N+1}$ ，包括正向和反向迁移

# 传统持续学习

离线持续学习

1) 任务增量学习; 2) 类别增量学习; 3) 领域增量学习

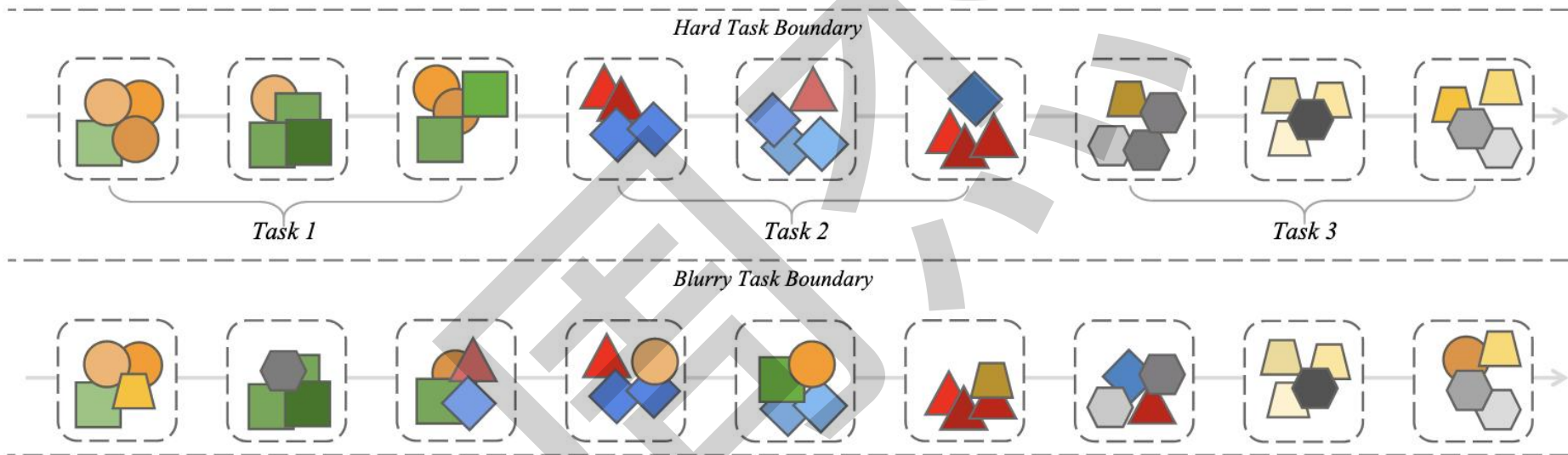




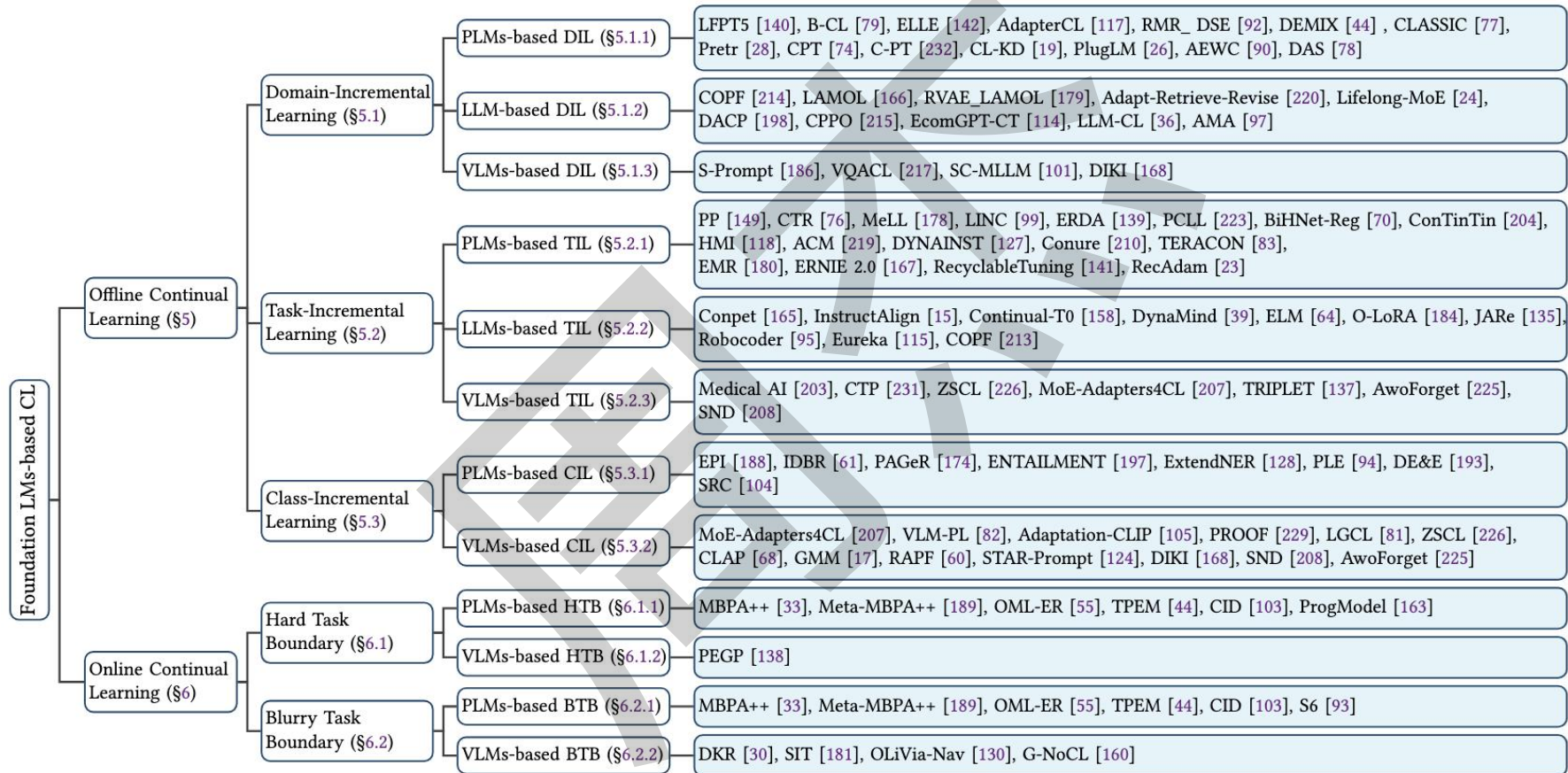
# 开放持续学习

## 在线持续学习

1) 固定任务边界; 2) 模糊任务边界



# 基于大模型的持续学习综述





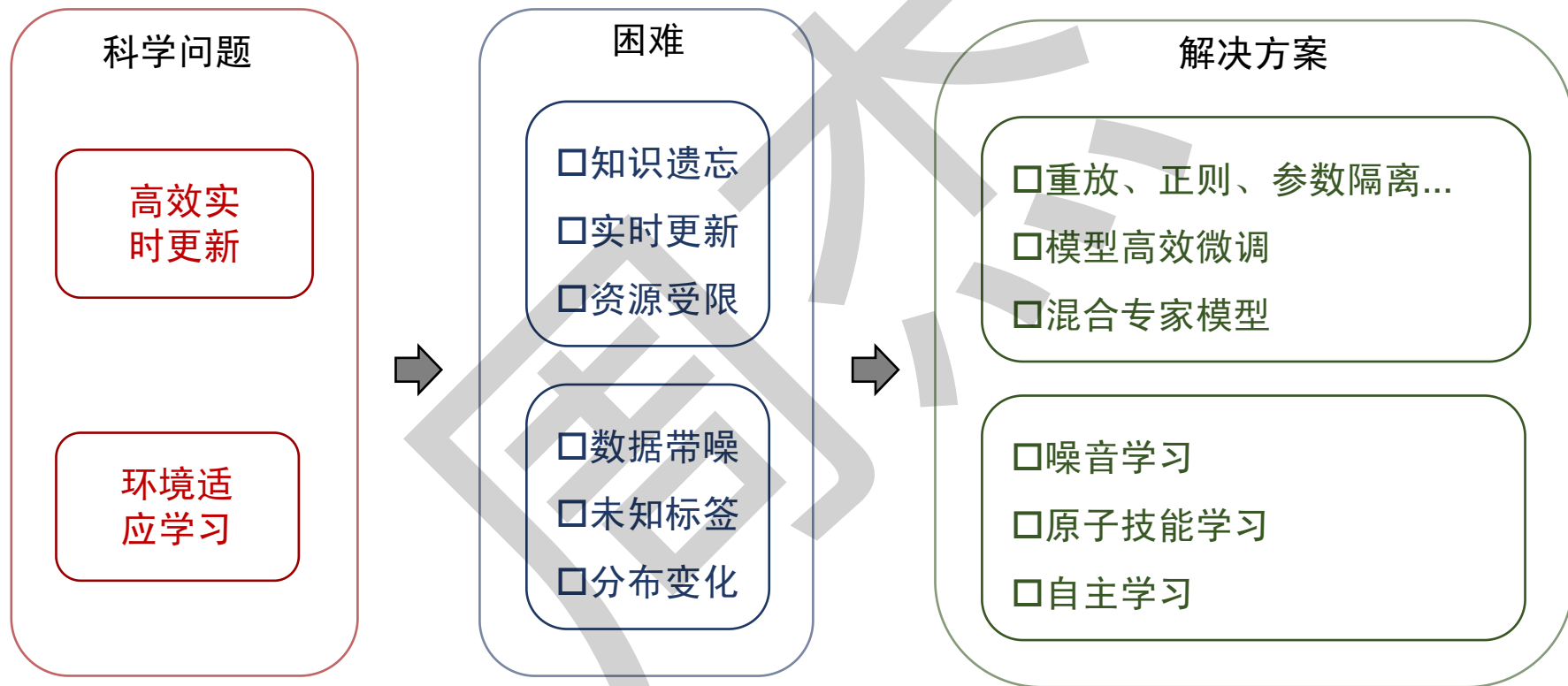
華東師範大學  
EAST CHINA NORMAL  
UNIVERSITY

## 目录 | CONTENT

- 持续学习背景介绍
- 开放环境大模型持续学习
- 持续学习发展趋势



# 开放环境大模型持续学习挑战

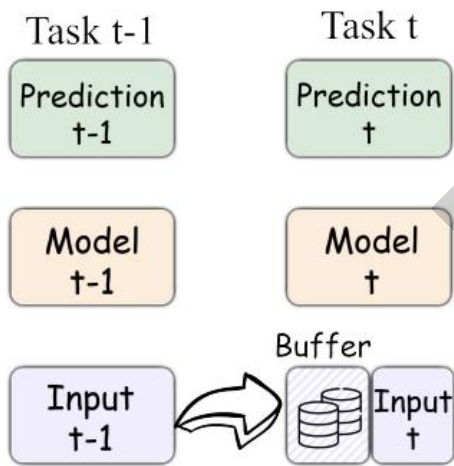




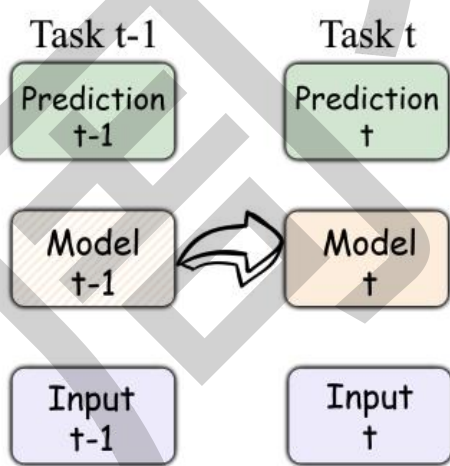
# 持续学习中遗忘问题

## 常见的三种持续学习方案

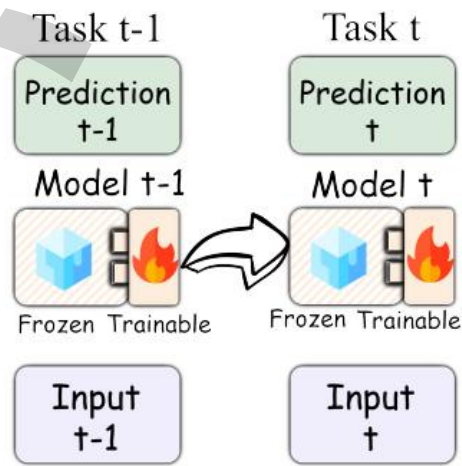
- 数据回放：采样少量历史任务数据进行训练
- 基于正则：利用正则对参数进行约束，不偏离太远
- 基于架构（如：参数隔离）：不同任务之间参数独立



(a) Replay-Based



(b) Regularization-Based

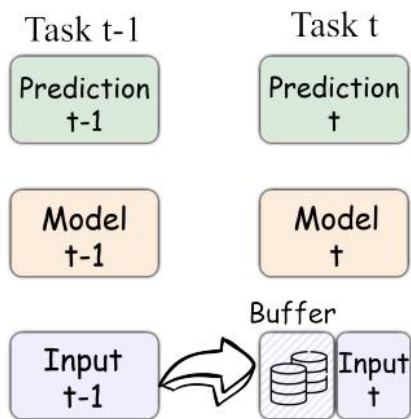


(c) Architecture-Based

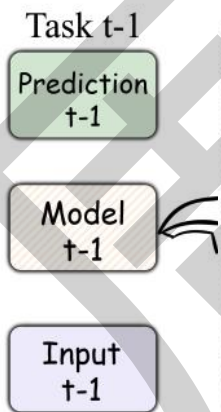
# 持续学习中遗忘问题

## 常见的三种持续学习方案

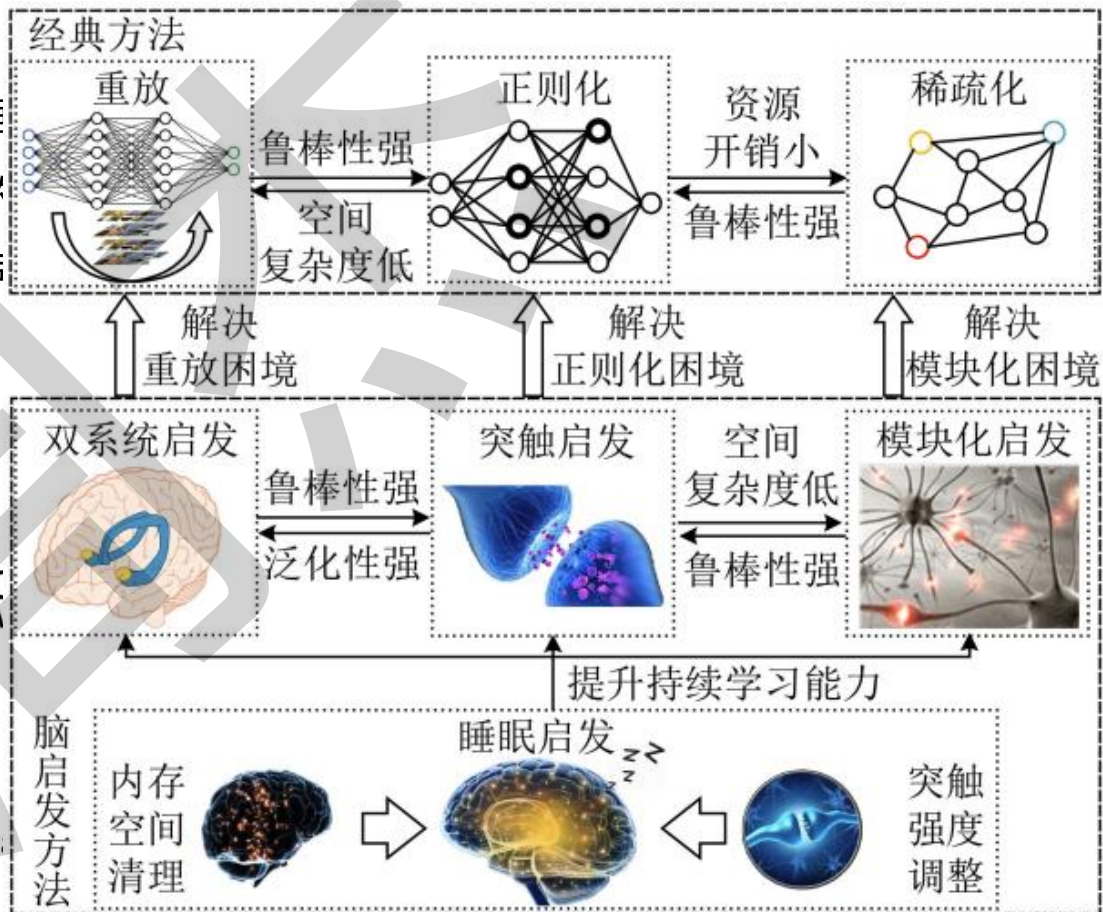
- 数据回放：采样少量历史任务数据
- 基于正则：利用正则对参数进行约束
- 基于架构（如：参数隔离）：不同



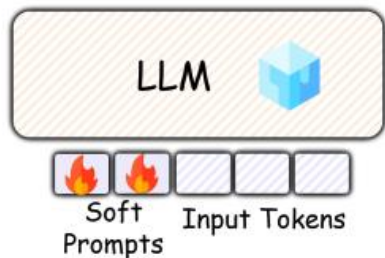
(a) Replay-Based



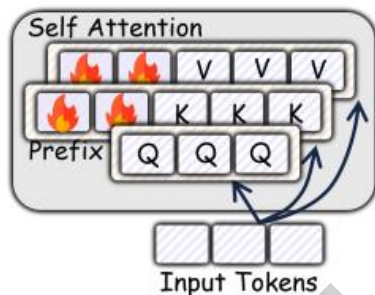
(b) Regularization



# 基于高效微调的持续学习



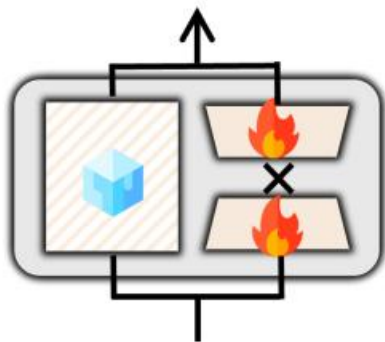
(a) Prompt Tuning



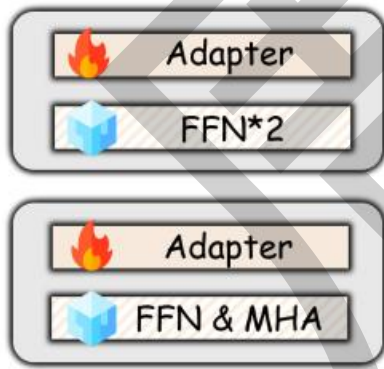
(b) Prefix Tuning

传统模型主要采用额外添加参数的方式

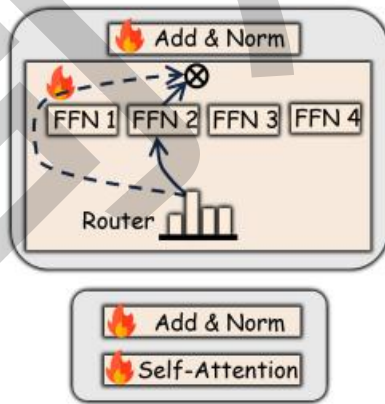
如何精准定位参数进行修改？



(c) LoRA

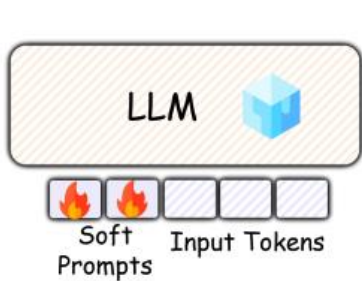


(d) Adapters

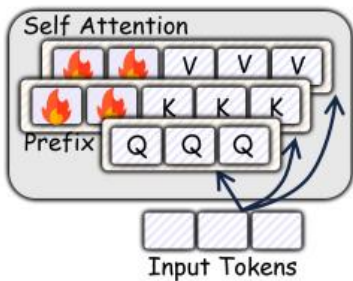


(e) Mixture of Experts

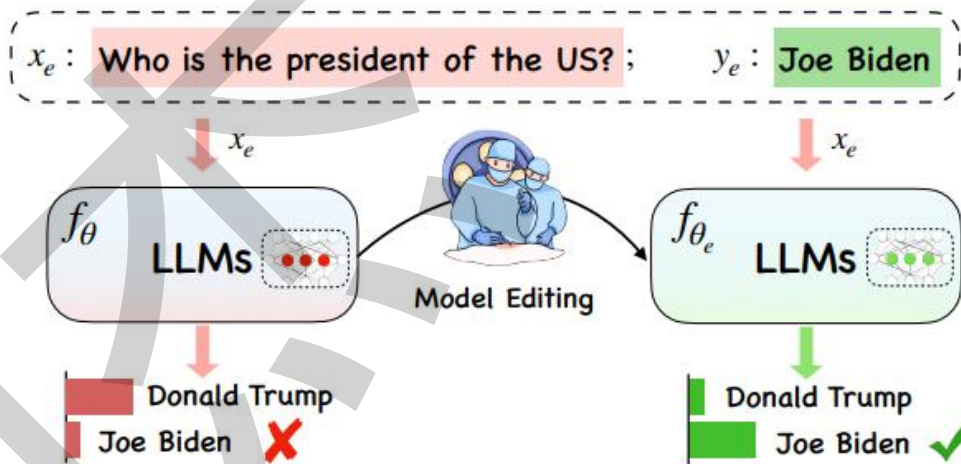
# 基于高效微调的持续学习



(a) Prompt Tuning



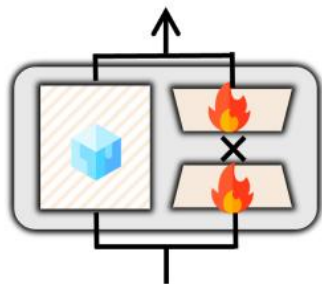
(b) Prefix Tuning



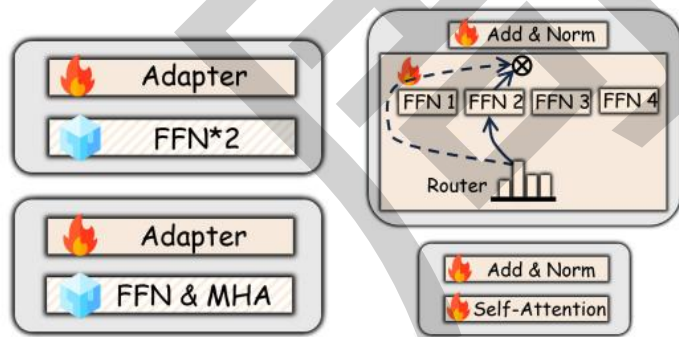
知识编辑

传统模型主要采用额外添加参数的方式

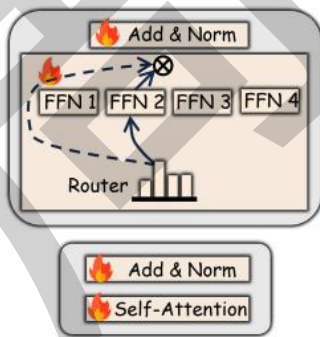
如何精准定位参数进行修改？



(c) LoRA



(d) Adapters



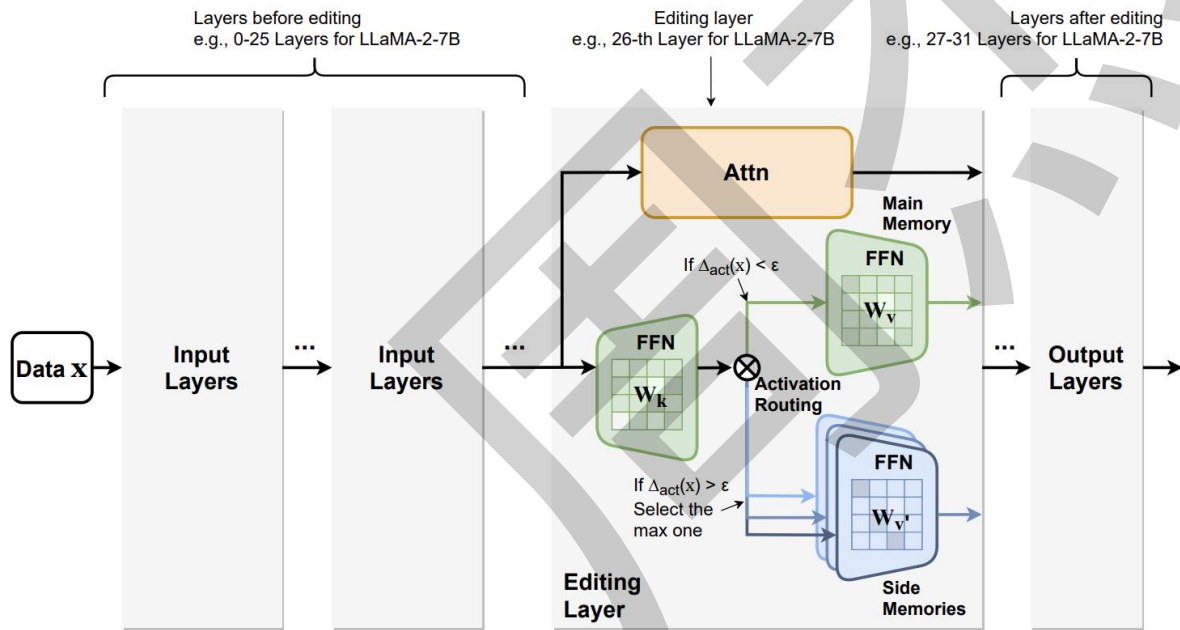
(e) Mixture of Experts



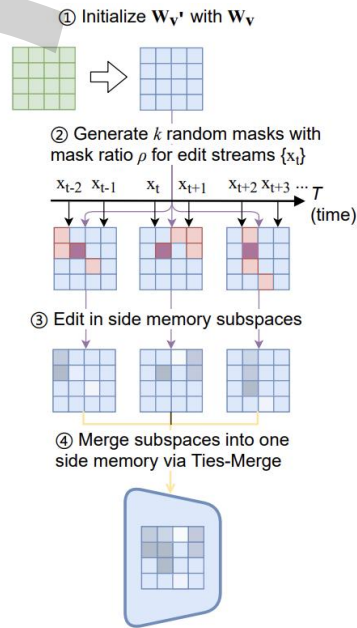
# 基于高效微调的持续学习

持续知识编辑算法，包含主要记忆和边缘记忆

- 编辑边缘记忆中的知识，通过路由器来分配记忆。
- 不同边缘记忆编辑不同的参数子空间，从而在没有冲突的情况下合并到共享记忆中。



(a) Workflow Overview with Knowledge Routing



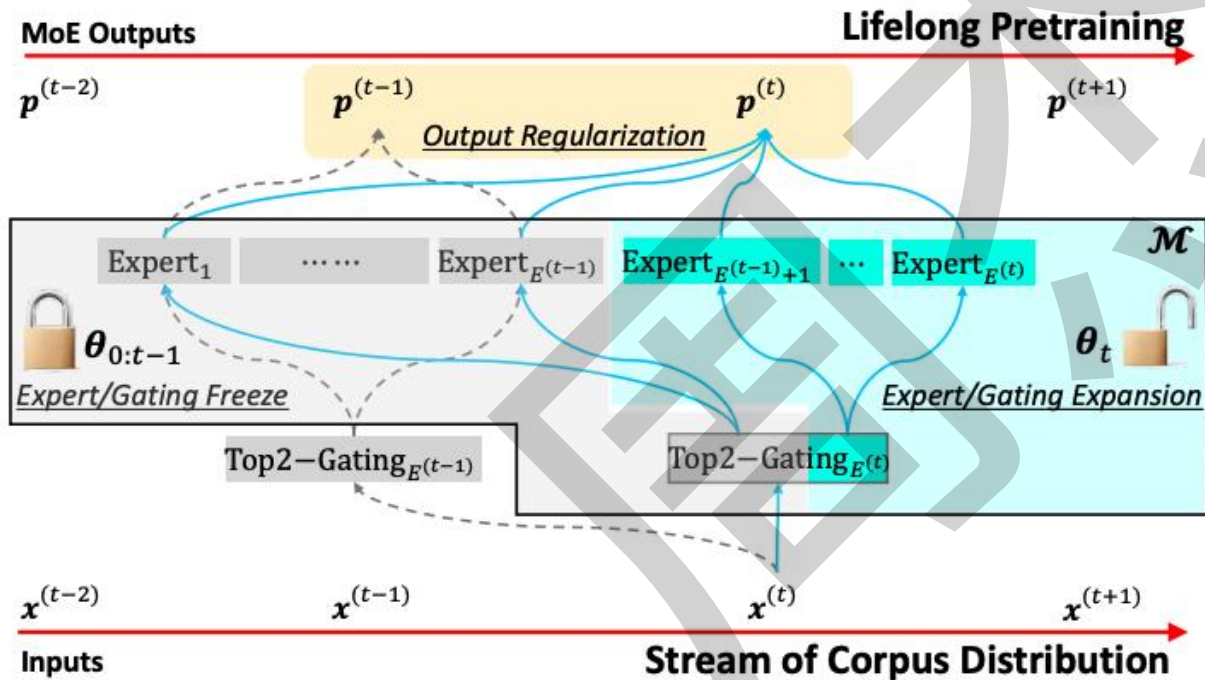
(b) Knowledge Sharding and Merging



# 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

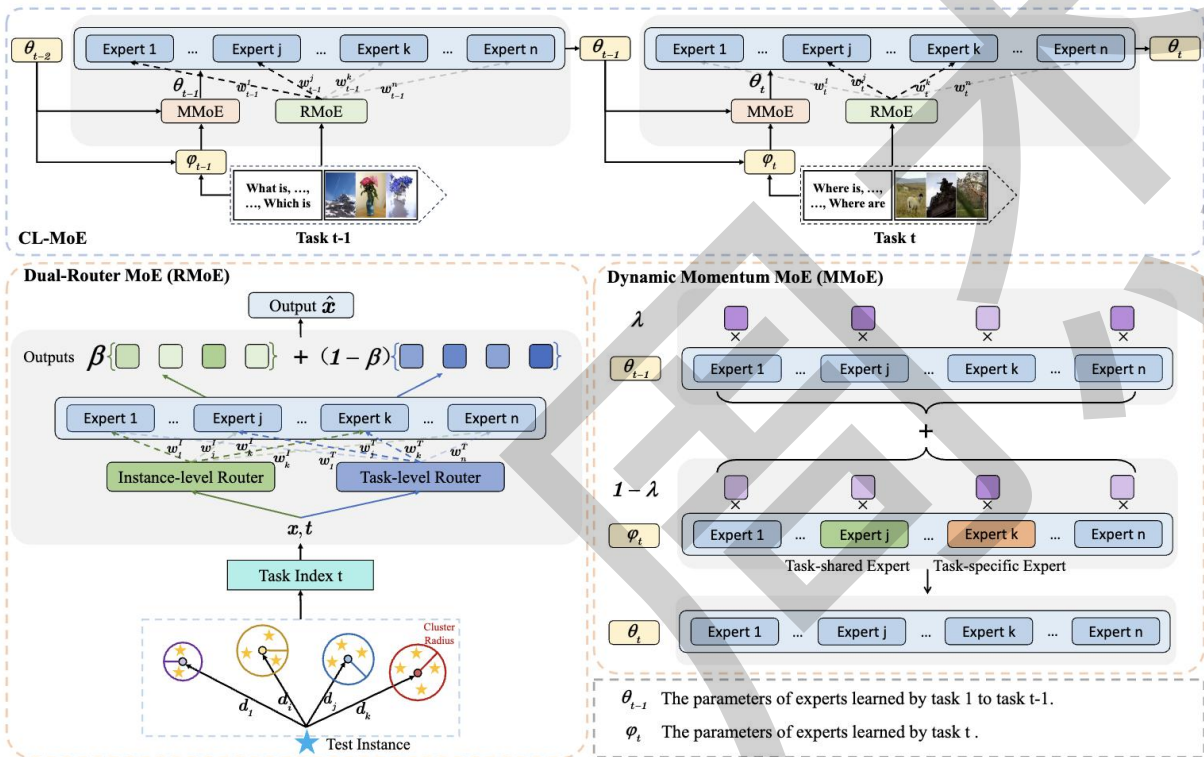
□ 原来的专家和 Gating 不更新

□ 隐式正则

□ 利用蒸馏的方式进行正则优化, 防止遗忘问题

# 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

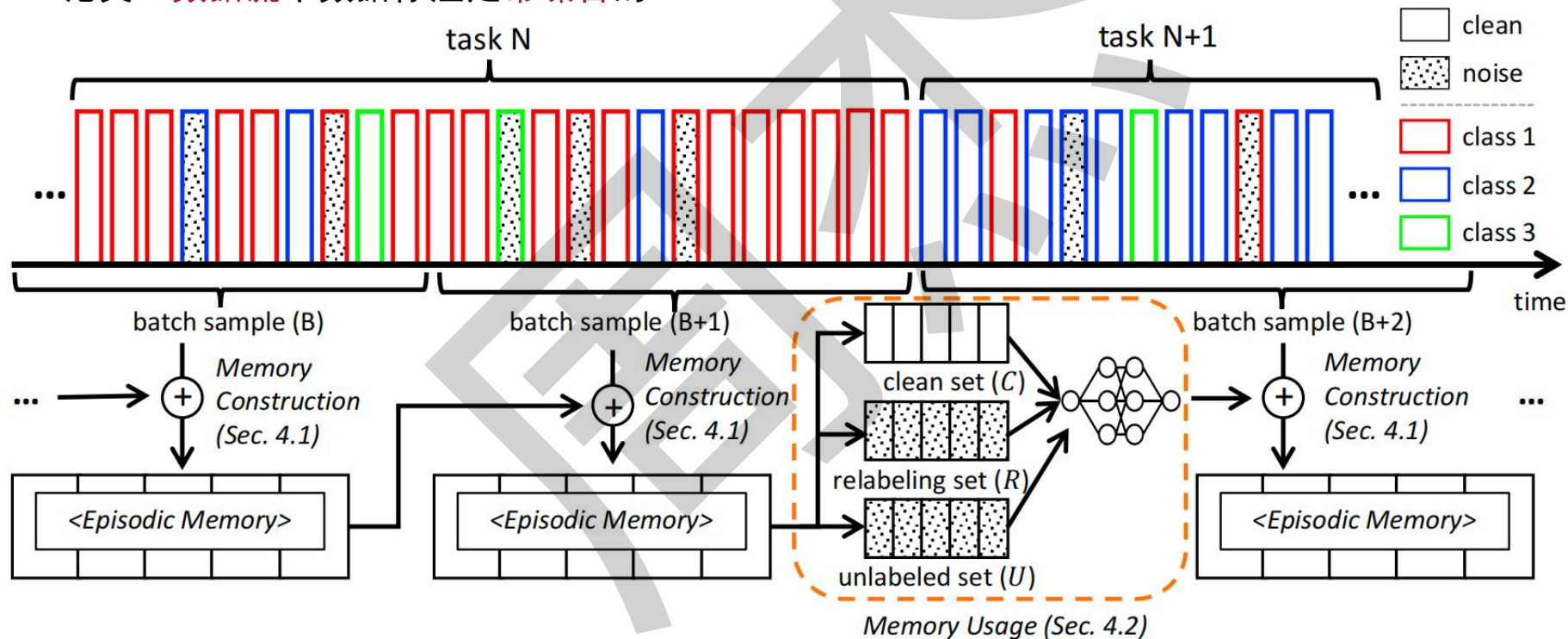
□ 一个关注**任务级别**全局信息

□ 动态动量更新

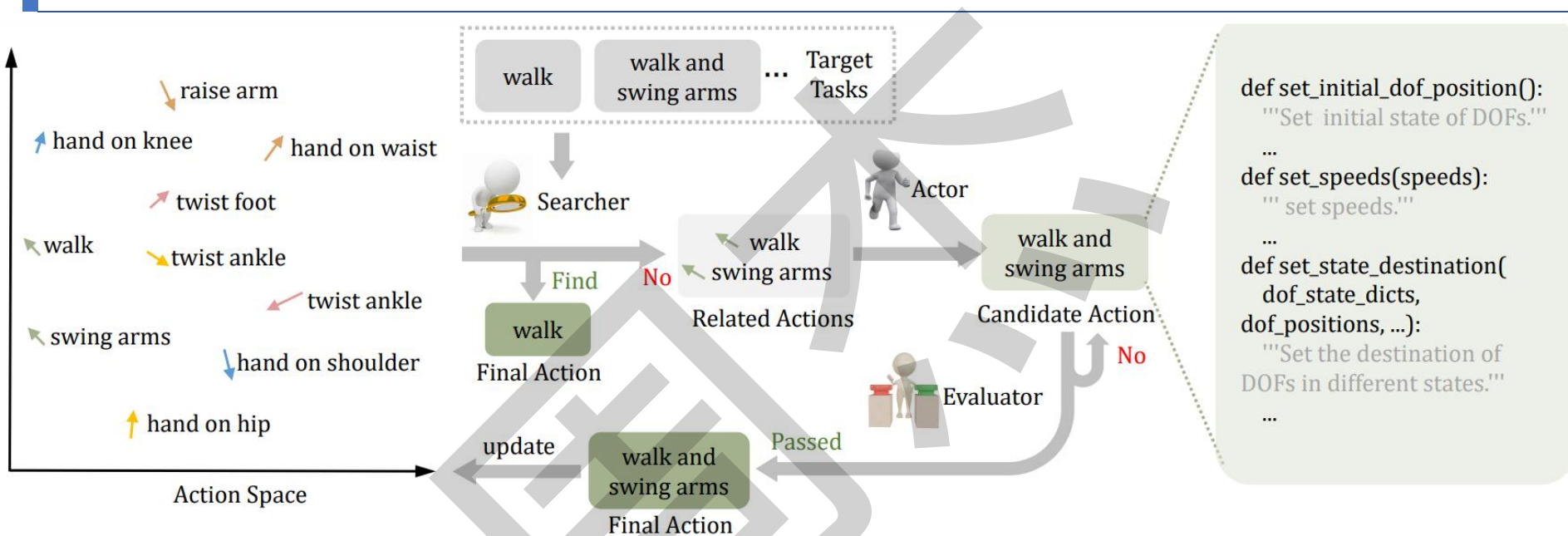
□ 对于任务**共享**专家和任务**特定**专家采用不同的更新方式

# 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



# 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

# 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

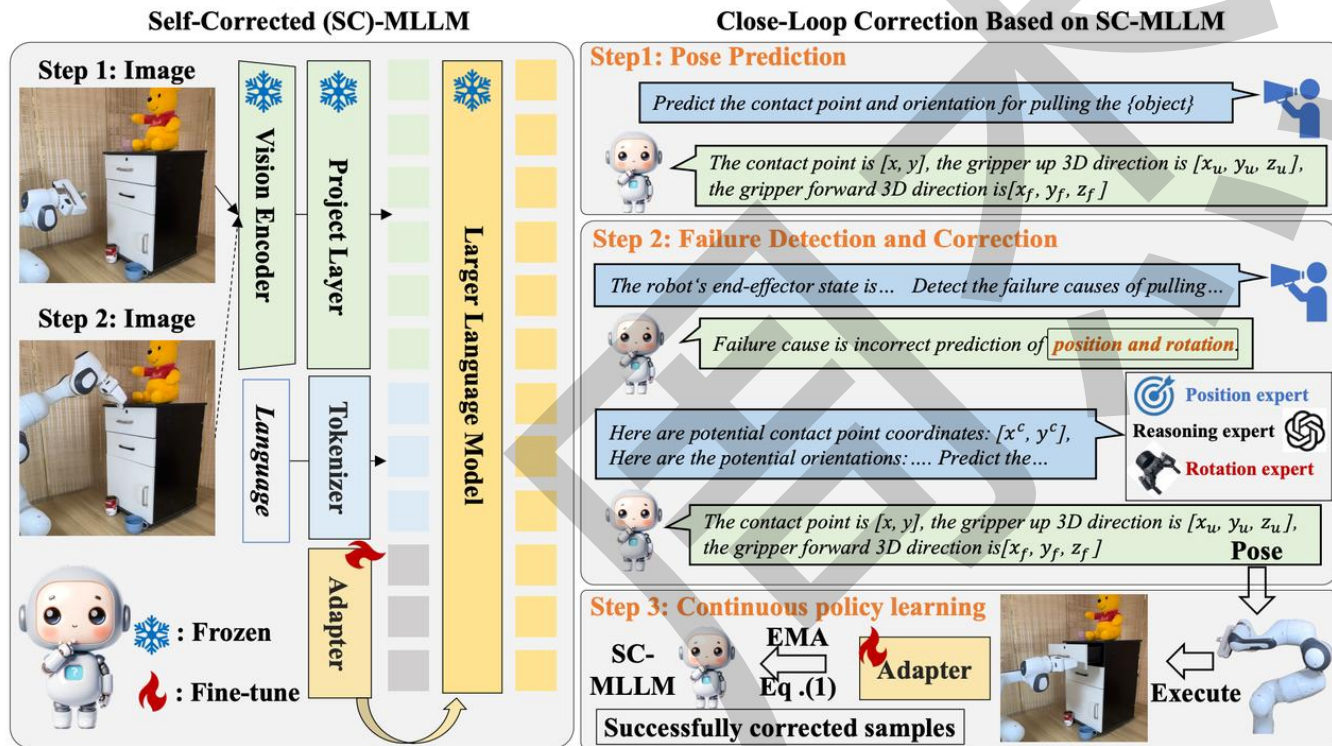
犯一次错是无知，  
犯两次错是愚蠢。

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







華東師範大學  
EAST CHINA NORMAL  
UNIVERSITY

## 目录 | CONTENT

- 持续学习背景介绍
- 大模型持续学习
- 持续学习发展趋势



# 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动力
- **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

# 可持续学习的个性化大模型

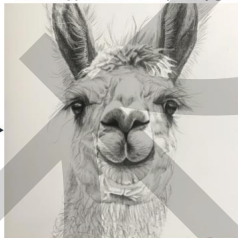
LLaMA V1.0



LLaMA V1.1



LLaMA V1.5



LLaMA V2.0



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

白天对话

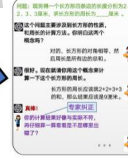
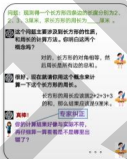
晚上进化

白天对话

晚上进化

白天对话

晚上进化



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”

谢谢!