CAAI-蚂蚁科研基金（AGI 专项）
申报课题介绍

目录：

5. 高性能 Agentic RL 系统研究
6. 面向大规模后训练场景下的训推一体引擎加速

（一） AGI 数据与评测

1. AIGC 视频评测大模型

课题背景：

随着生成式人工智能技术加速渗透视频生产领域，文本/图像到视频的生成效率已实现突破性进展，但评测体系的滞后性问题日益凸显。当前行业普遍面临三重挑战：其一，传统视频质量评估（如 PSNR、SSIM、MOS 等）局限于分辨率、噪声和视频无参考分检测，无法有效识别 AIGC 视频特有的主体形变（如肢体逻辑错误）、背景时空断裂（如物体突然消失/重现）、音画异步（如唇动与语音偏移）等新型缺陷；其二，人工标注成本与视频生成效率严重失衡，头部平台日均视频生成量超百万条,但现有自动化评测在语义对齐度(如提示词指令遵循偏差)、视频时序一致性（如主体、背景突变）、声画合理性（如声画不同步、不匹配）等多个维度缺乏量化标准；其三，企业图生视频服务面临生成效果不稳定问题，例如同一输入指令在不同模型迭代版本中生成视频的图/提示词遵循度、美学质量等可能存在巨大差异。这些矛盾不仅制约了 AIGC 视频的商业化落地，更阻碍了多模态生成技术的迭代闭环。建立覆盖“生成-评测-反馈”全链路、融合“传统视频、AIGC 视频以及指令生成视频共性与特性”的全维度评测体系，已成为产业界与学术界的共同诉求。

目标和产出：

1.开源 AIGC 视频评测大模型/系统一套，评测指标可量化，覆盖至少 10 种以上关键评测维度，评测系统中自研模型不少于 3 个
2.申请至少 2 项发明专利
3.发表领域内顶级会议或期刊高质量论文至少 1 篇

2. 大模型高效数据蒸馏方法

课题背景：

语料合成是 AGI 的关键能力之一，在蚂蚁已引入的大量优质文本语料基础上，利用先进语言大模型的优势能力，实现高质量、稀缺领域知识的快速获取和提纯，成为进一步优化训练语料、提升模型训练效果的重要选择之一。在对业界新模型的知识蒸馏的过程中，种子 query 的质量和多样性作为影响合成语料训练效果的核心因素，依然存在提升空间：
1. 缺乏能代表 AGI 各领域和用户真实需求的优质 query 集合，在更高难度或更精细领域的种子生产或筛选上，需要引入更多业界前沿的 query 生产方式，沉淀和生产各类专业领域高质量、高难度、多样性的百万级专业种子 query，通过优质 Query 进一步提升知识蒸馏效果上限。
2. 模型知识蒸馏技术的全面性提升，通过学术背景的人员参与，对高难度推理、模型自蒸馏、

分阶段蒸馏、开源模型白盒蒸馏等前沿蒸馏技术的进一步研究，沉淀一套业界先进高效知识蒸馏方案，尽可能覆盖业界新模型的优质知识数据，充分利用先进模型的优势能力，来提升蚂蚁模型性能。
3. 完善蒸馏效果的评测体系，需要沉淀一套针对不同领域合成数据的评估 benchmark，能根据知识蒸馏结果快速验证业界新模型的优势能力，高效产出分析结果和评估报告，为蚂蚁模型各个训练阶段的语料合成场景提供指导，供给优质的知识蒸馏合成语料，提高蚂蚁大模型训练效果。

目标和产出：

1.交付 100w 精细领域的高质量、高难度、高多样性的优质 Query。
2.产出一套能有效评估数学推理、代码、STEM 等领域知识蒸馏效果的 benchmark，高效发现和验证业界新模型的先进能力，为蚂蚁在模型训练各个阶段的语料合成场景提供参考。
3.产出能有效提升蚂蚁知识蒸馏效率效果的技术方案（包括不限于：数学、代码、STEM、逻辑等领域 query 的评估筛选、增强、生成能力、前沿的模型知识蒸馏技术方案）。

3. 大模型动态评测和反污染检测方法

课题背景：

静态榜单的局限性与“刷榜”问题：当前大模型的评测大多依赖于静态榜单，例如各类公开数据集和竞赛排行榜。然而，这些榜单题目常常由于其开源性或信息泄露，导致参与者能够轻易地通过构造相似题目或合成榜单题目等方式实现“刷榜”。这种现象使得静态榜单无法真实反映模型的能力进步，也难以对大模型的健康发展形成正向反馈，甚至可能误导研发方向。

模型训练数据污染的检测难题：另一个亟待解决的关键问题是模型训练数据是否被“污染”，即榜单数据或其高度相似的数据是否被不当纳入了模型训练。目前，业界缺乏直接且有效的手段来检测模型训练过程中是否存在这种数据“去污”（data contamination）或数据泄漏（data leakage）行为。这种不透明性不仅损害了评测的公平性，也使得我们无法准确评估模型的泛化能力和真实水平。

目标和产出：

1.不低于 5 套动态评测种子集。
2.一套评测集动态化算法包，为构建持续更新且难以预测的评测场景提供技术支撑，确保评测的公平性和有效性。
3.一套反污染检测算法集：集成多种先进的检测策略，对模型训练数据中是否存在榜单数据或高度相似数据的“污染”情况进行量化评估和有效判别，为大模型的可靠性评估提供坚实的技术保障。

（二） AGI 基础模型

1.基于代码图的仓库&跨仓库级别的代码理解与推理（强化学习）

课题背景：

在大型软件开发中，高效、准确地理解和操作仓库级乃至跨仓库级的代码至关重要。现有方法普遍将代码库视为非结构化的超长文本，导致两大核心瓶颈：（1）理解与推理低效：基于长文本的分析不仅计算开销巨大，且因缺乏代码内在的依赖和调用结构，导致推理结果不精准、上下文易丢失。（2）执行反馈昂贵：代码优化和 Agentic RL 等任务依赖的执行反馈，在仓库级反复进行会产生高昂的算力与时间成本，使得需要迭代试错的强化学习几乎不可行。

为突破上述瓶颈，本课题提出以代码图（Code Graph）作为代码库的结构化知识核心，将海量源码转化为包含函数调用、类继承、文件导入等关系的图谱。基于此，我们旨在研究：（1）高效的代码图推理技术：通过在代码图上进行精准的图遍历与剪枝，替代对超长文本的暴力分析。这不仅能极大降低计算复杂度，还能利用图的拓扑结构进行更深层次、更准确的逻辑推理和影响域分析。（2）低成本的图上强化学习框架：将高成本的物理代码执行反馈，部分转化为低成本的图上逻辑推演。通过在代码图上模拟变更的局部影响，快速评估动作（Action）的优劣，从而将强化学习的训练迭代速度提升数个数量级，显著降低资源消耗。

目标和产出：

1.技术产出：产出基于 Code Graph 的高质量模型和 agent，在仓库级别权威榜单 swe-bench 上带来 5%的提升，做到业界前三。
2.业务价值：帮助蚂蚁 AI coding 中的 new feature2code 相关业务指标提升 5%。
3.论文及专利：发表领域内顶级会议或期刊高质量论文至少 1 篇；申请至少 1 项专利。

2. 基于大语言模型的多轮多步工具调用技术研究

课题背景：

大语言模型通过调用外部工具扩展能力、执行人类指令、完成复杂任务，这对于弥合大模型理解与行动之间鸿沟具有重要意义。然而，当前大模型在工具学习（Tool Learning）任务上仍面临诸多挑战：1）用户意图与工具特性理解不足； 2）复杂任务的规划与执行能力有限；3）工具使用的可解释性与可控性缺乏透明度；4）工具学习的数据构造与模型训练效率低；等等。这些不足严重限制大模型在人类真实复杂任务上的实用性，亟需理论突破与技术框架创新。

建议研究方向：1）面向真实世界的复杂任务理解、规划与执行：探索大模型内省推理与任务规划能力的融合，解决大模型在任务执行前对用户意图和工具特性的精细化理解不足；任务执行中长程规划、决策和执行能力弱；任务执行后决策逻辑追踪分析欠缺。
2）面向开放、动态环境的持续学习：构建基于强化学习等技术的新颖算法和技术框架，实现大模型与环境直接交互，探索大模型外省推理与任务规划能力的融合，形成模型自进化能力。

目标和产出：

1.模型赋能：技术应用于大模型的行动力提升，BFCL-V3 基准评测的总排名和多轮工具调排名进入 Top5。
2.技术产出：提出新颖 Tool Learning 算法，形成工具学习框架，沉淀高质量工具调用数据集（具备场景覆盖完备、贴近真实世界、意图-行为精准对齐、上下文依赖完整、标注体系规范等特性），构建开源 SOTA 模型。
3.论文及专利：发表领域内顶级会议或期刊高质量论文 2 篇；申请 2 项专利。

3. 基于视觉语言大模型的下一代 OCR 技术

课题背景：

OCR 和广义的屏幕理解相关算法是通用文档理解、复杂表格结构化、图文混排识别等任务中不可缺少的功能。对于过往的基于检测+识别方案的 OCR 而言，由于无法识别语义，其结构化输出强烈依赖人工预先标注，包括人工标注，和人工后处理规则的编写。标注和规则编写工作复杂，效率较低，且容易出现版本不匹配、人工疏漏导致的结果错误，无法满足业务快速迭代发展时，对识别准确性和语义理解能力的要求。将多模态方案与当前业务场景相结合，可有效提升算法准确率，为业务带来更多智能运营和终端洞察能力。大型多模态模型由于其检测识别理解一体化的特性，一个模型即可端到端地对复杂场景（图文混排、复杂表格、多栏文档等）实现内容理解和结构化输出，大幅降低累积误差，同时其端到端的特性可有效减少对后处理脚本的依赖，需要调整的超参数也大幅度减少，无论是前期标注，中期算法开发，还是后期后处理规则编写，都能大幅减少相关人力成本，提升相关功能迭代效率。

目标和产出：

1.模型赋能：技术应用于表格、文档、照片等类型图片理解任务中，json 语法正确率高于 99.8%，提高现有业务线上服务质量，驱动产品应用创新。
2.技术产出：提出具有创新性的模型架构或方法论，在 QA 类数据集（如 OCRBench），综合类 OCR 评测数据集达到 SOTA 水平（如 CC-OCR bench、omnidocbench、截止 25 年 7 月，MonkeyOCR3B 在 omnidocbench 中英文的指标分别为 0.277/0.140，GPT4o 为 0.399/0.233，归一化编辑距离越低越好）。对于不涉及敏感、隐私数据的部分，开源相关代码，提高行业内技术影响力。
3.论文及专利：发表领域内顶级会议或期刊高质量论文 2 篇；申请 2 项专利。

4. 面向高效训练与推理的混合专家模型（MoE）算法设计与优化

课题背景：

人工通用智能（AGI）领域的突破高度依赖于基础模型的性能提升，而混合专家（Mixture of Experts）模型因其能够在不过度增加计算成本的前提下有效提升模型容量，已成为当前预训练大模型架构设计的研究热点。然而，现有的 MoE 模型算法在专家路由策略、专家间负载均衡性以及计算资源高效利用等方面仍存在诸多挑战，导致在实际工程应用中存在训练效率低、模型稳定性差、推理成本高等问题。这些算法层面的瓶颈直接制约了大模型的进一步规模化和落地应用。

目标和产出：

1. 新的模型架构能有效应用集成到百灵下一代模型；
2. 新的架构效果和成本均优于现有业界主流模型（如 deepseek-v3）；
3. 发表领域内顶级会议或期刊高质量论文 1 篇；
4. 申请相关算法设计的技术专利至少 1 项。

5. 面向前沿大模型架构的算法和软件联合加速技术研究

课题背景：

AGI 进入下半场之后，新的大模型架构创新层出不穷，包括但是不限于：
1）LLM 中支持递归甚至动态递归的、支持不同 MoE 设计的新型网络；2）支持多模态甚至全模态理解、全模态理解生成一体等新的架构；3）多模态生成场景，从多步迭代 diffusion 到更少步骤甚至单步迭代的 diffusion；4）全新的基于 diffusion 的 LLM5）面向 testing-time-scaling 而言的一些新的架构创新 6）带有 memory 机制的大模型架构创新 7）线性注意力和混合架构的推理加速。

这些新的架构创新都给模型推理加速带来了新的挑战。在这些新的场景，可以从事包括但不限于如下的研究：1）专用架构的细粒度算法或者软件实现联合优化 2）从现有架构，无需训练和少数据训练，迁移到其他高效架构 3）通过 training-free 或者 few-sample calibration 优化模型信息流，达到提高模型效果甚至推理效率的目标 4）通过理论指导实践，实现推理加速。

目标和产出：

1.用于新型架构推理加速的代码库、框架或者完整 toolkit；
2.在某一具体方向相比 SOTA 实现显著加速（30+%），并产出实践报告；
3.申请至少 1 项专利；
4.发表领域内顶级会议或期刊高质量论文至少 1 篇；

6. 原生多模态大模型交互体验升级优化

课题背景：

现有多模态大模型在诸多理解任务上表现出了很强的能力，使得模型具备了很强的工具属性、可以帮助用户解决具体的问题，这种较高的“实用价值”我们可以认为模型具备了很高的“智商”；但在真实生产生活场景下：
1）原有交互范式缺陷：受到地域、文化以及个体差异等影响，用户在与大模型交互过程中往往输入的指令表达风格千差万别，甚至可能表达简略模糊残缺、甚至包含语病与错误，也可能是令晦涩复杂难懂的，而模型在此类场景下回答质量往往会下降、甚至无法正确 get 用户的意图出现“已读乱回”的情况，影响模型的“情绪价值与实用价值”。
2）新交互范式升级：现有大模型在与用户交互过程中更多的是“被动响应”模式，即用户问什么、模型就回答什么，而真实人类交互过程是包含“主动响应”模式的，即交互过程双方有问有答：在交互过程中，模型如果可以做到的“主动”揣测、分析、挖掘用户在指令意图表层含义之上的想法，站在用户的角度为用户着想、做到对用户的深层意图的提前感知与响应，甚至可以对整个交互过程提供一定的引导性，从而模型就可以为用户提供更积极主动的“情绪价值与实用价值”。

为此，本项目旨在探索提升多模态模型提升“情商”方面的智能上限的方法，做到“更懂用户”，进而提升模型的对话的体验。具体来说主要解决以下问题：
1）体验 benchmark 设计：探索多模态细体验的 benchmark 评估方案，为多模态模型的综合体

验性能优化提供科学明确的牵引。
2）体验对齐优化方案设计：探索多模态细体验对齐优化方案，提升多模态模型的体验性能。

目标和产出：

1.业务赋能：在视频聊天机器人、具身智能机器人等新业务场景落地。
2.技术指标：在多模态交互场景下体验性能达到“多模态大模型业界 SOTA”（公开榜单（比如 Chatbot Arena / OC arena / superclueV）。
3.论文及专利：发表领域内顶级会议或期刊高质量论文 1-2 篇；申请 2 项专利。

7. 面向 Web 的推理时交互的多模态 UI 智能体关键技术研究

课题背景：

GUl-Agent 是最近两年在大模型和 AIAgent 迅猛发展的浪潮下的热点方向。目前，工业界和学术界内都涌现出了一批开源项目和产品，并引起了多方的广泛关注，包括微软的 GUI-Actor，字节的 UI-TARS 等等。这个方向具有极高的应用价值，在新人机交互、流程自动化等领域。

然而，目前的方案还存在不少问题，如页面元素漏检、tokens 消耗过多、未见过的站点通过率较低等，这些问题和挑战阻碍了 GUl-Agent 在商业化场景的应用和落地。本项目希望通过前沿的多模态、强化学习等领域的前沿技术解决现有 GUI-Agent 的不足。使得这项技术能够更好地落地到数科的业务和商业化产品。

我们期待的研究尝试，围绕在“面向 web 的推理时交互” （Interaction with Web at Inference Time）。传统的基于静态 DOM 解析与网页截图的网页理解方法，在处理现代复杂、动态的 Web 应用时面临固有局限性。大量关键信息与交互元素并非在初始加载时即显现，而是依赖于用户的特定操作（如鼠标悬停、滚动、点击等）才被激活或展示。为克服此瓶颈，我们希望构建“面向 Web 的推理时交互”框架。该框架旨在赋予 LLM 智能体在推理过程中主动运用前端交互工具的能力，从而增强动态复杂网页环境中收集无法由 DOM 解析和单次页面截图充分展现的信息，从而提高单次 action 以及 task 的执行成功——这显然是 GUI 智能体产生实际价值必须跨越的门槛。

目标和产出：

1.发表领域内顶级会议或期刊高质量论文 2 篇；
2.代码移交及落地验证，在 Webvoyager 等公开数据集、蚂蚁内部构造业务数据集上，较 browser-use 等开源 SOTA 方案的效果提升 10%

8. 面向复杂任务的多模态统一推理框架

课题背景：

近年来，多模态大模型（MLLMs）在整合文本、图像、音视频等多模态信息方面取得了革命性进展，展现了强大的跨模态感知与初步理解能力。然而，当面临数学逻辑求解、精细空间关系判定、以及图形用户界面（GUI）的程序化交互等需要深度、结构化推理的复杂任务

时，现有模型普遍暴露出能力瓶颈：1）显式推理模式的低效性，制约了其有效推理深度；2）泛化能力受限，知识难以跨模态、跨任务迁移；3）信息解析粒度失配，无法根据任务需求动态调整理解深度。这些局限性严重阻碍了 MLLMs 从“感知”向“认知”的跨越。本项目旨在直面这一挑战，通过聚合数学、空间、GUI 三大关键领域的推理能力，构建一个统一的深度推理框架，并探索其背后的可泛化机制，为实现更高阶的多模态人工智能奠定基础。

目标和产出：

1.在隐空间推理/动态感知/数理逻辑推理方向产出 1 篇高质量论文，多模态推理指标达到业界前三；在空间推理/GUI 推理方向产出 1 篇高质量论文，多模空间推理/GUI 推理指标达到业界领先。
2.支撑多模复杂推理模型在探一下、智能交互等业务场景中达到开源模型 SOTA，并落地应用。

9. 面向新一代硬件架构的大语言模型大规模分布式训练提效方法研究

课题背景：

以 chatgpt、deepseekV3 为代表的语言大模型的预训练通常需要消耗巨量的计算资源，如何充分利用计算资源、提升语言大模型训练效率成为一个急需解决的问题。deepseekV3 通过良好的算法-工程联合设计，在特定加速卡上达成了非常高效的训练。

近年来，GPU 及其他厂商的 AI 加速芯片也在持续迭代中，新的硬件架构带来了新的特性，例如以华为和 NV 为代表的超节点技术逐渐成熟，超节点利用超高速全互联使得数百张 GPU 卡拥有跟单机多卡相似的互联带宽。

本项目旨在研究下一阶段面向新一代硬件架构背景下各类提升语言大模型训练效率的方法，包括但不限于：1）通过算法工程联合设计，使得大模型架构能够完全适配硬件特性及训练引擎的软件特性，达成最高的训练效率；2）构建系统化的性能预估和落地体系，更合理引入业界最新的性能优化技术；3）针对蚂蚁大模型的业务和技术特点，设计和落地专用的性能优化技术。

目标和产出：

1.利用新硬件架构特性，在典型开源模型和蚂蚁自研模型上进行优化，使训练性能(按 MFU 计算）较开源方案提升 20%以上。沉淀可在蚂蚁落地的方案的源码、设计文档、使用文档，帮助蚂蚁语言大模型训练提效。
2.上述目标实现后沉淀可在蚂蚁落地的方案的源码、设计文档、使用文档，帮助蚂蚁语言大模型训练提效。
3. 发表领域内顶级会议或期刊高质量论文至少 1 篇；申请至少 1 项专利。

10. 基于原生多模态大模型的像素级细粒度理解技术研究

课题背景：

让模型能够在复杂的现实世界场景中同时感知、理解并推理多模态/全模态输入（如文本、视频和音频），仍是人工智能领域的一个长期目标。近年来，全模态预训练和指令微调

方面的进展催生了全模态模型，使我们更接近这一目标。尽管取得了这些进展，当前的多模态/全模态模型在两个关键领域仍存在明显局限：（1）对视频和音频中复杂时间序列的长时程推理能力；（2）像素级的细粒度空间理解能力。

目标和产出：

1.技术突破：解决长视频理解场景难题，
2.产业落地：为智能 AI 助手、小时级长视频理解等场景提供可落地的解决方案，显著提升人机交互的自然度与用户体验；
3.学术研究：首提基于 GRPO 的时序感知选帧框架，解决动态信息建模难题。推动多模态理解领域创新，形成具有国际影响力的科研成果。
4.论文及专利：发表领域内顶级会议或期刊高质量论文 1-2 篇；申请 2 项专利。

11. 面向全域复杂文档的多模态大模型解析技术研究

课题背景：

面向全域复杂文档的多模态大模型解析技术，核心在于通过融合文本、图像、表格、图表等多种模态信息,实现对复杂文档的深度理解与高效处理。复杂文档的多模态解析技术通过“感知-认知”融合，正在重塑文档处理的范式，为通用人工智能的发展提供关键支撑：
1）突破单模态局限，实现跨模态协同解析：多模态大模型通过融合文本、图像、表格、图表、公式等多类型数据，解决了传统单模态方法（如纯 OCR 或纯文本分析）的信息孤岛问题，显著提升信息提取的完整性与准确性
2）增强复杂结构理解与动态优化能力：面对多栏排版、嵌套表格、手写注释等复杂文档结构，多模态模型通过端到端训练自适应学习布局特征。此外，模型还能动态分配资源，优先处理高信息密度区域（如财务报表中的关键数据）。
3）推动专业领域智能化升级：在医疗、法律、金融等领域，多模态解析技术可深度挖掘文档中的隐含信息，这种跨模态推理能力成为行业智能化的核心驱动力。
4）应对小样本与长文档挑战：针对法律、古籍等数据稀缺领域，多模态大模型通过预训练结合领域适配器（Adapter），在有限数据下仍能保持泛化能力 。同时，技术趋势表明，分块处理与全局注意力机制的结合正在突破百页级长文档的建模瓶颈 。
5）加速多模态知识库构建与应用：文档解析与向量化技术的结合，不仅提升了数据处理效率，还优化了多模态知识库的构建质量。

目标和产出：

复杂文档图像，包括扫描文档、表格、表单、发票、合同、手写笔记等多个形式，是金融、医疗、教育等场景的关键依赖，也是专业领域高密度知识数据的承载体，是多模态大模型的重要应用场景，同时也是高质量数据的重要来源。关键目标如下：
1.攻克表格、公式、图表等复杂结构的跨模态理解难题，关键技术能够集成至百灵多模态大模型；
2.在全域复杂文档评测集上，相对与百灵 Baseline 提升 10%；
3.论文及专利：发表领域内顶级会议或期刊高质量论文 1 篇；申请 2 项专利。

12. 多模态生成与理解任务的协同优化统一模型研究

课题背景：

多模态生成与理解任务的协同优化是指在处理包含多种模态（如文本、图像、音频、视频等）的数据时，通过同时优化生成任务和理解任务 ，提升模型整体性能的一种方法。这种协同优化不仅有助于模型更好地理解多模态数据，还能增强其生成能力，从而实现更自然、智能的人机交互。理解与生成的统一不仅是技术趋势，更是实现“感知-认知”融合的核心。它通过架构创新与训练策略优化，解决了模态对齐、数据稀缺等挑战，为未来人机交互和 AGI 发展奠定了基础。

目标和产出：

推动多模态生成与理解二者走向统一架构、共享表示、双向流动，是多模态大模型进一步突破的关键技术，是实现全模态自由交互重要依赖。关键目标如下：
1.探索图像理解与生成的共享表征算法，建立理解与生成统一 Tokenizor 新方案，理解与生成侧效果相对百灵 Baseline 提升 5%；
2.设计 3 种以上理解与生成相互促进的新型训练任务；
3.论文及专利：发表领域内顶级会议或期刊高质量论文 1 篇；申请 2 项专利。

13. 基于长上下文高效注意力机制的 Transformer 研究和设计

课题背景：

Transformer 架构的演进与瓶颈：自 2017 年 Transformer 架构提出以来，其核心的自注意力机制彻底改变了深度学习领域的发展轨迹。通过摒弃循环结构并引入全局依赖建模能力，Transformer 在机器翻译任务中首次超越 RNN 模型，随后催生了 GPT 等划时代的大语言模型（LLMs)。大规模语言模型（LLMs）的出现，极大的改变了自然语言处理领域的研究范式与面貌，并在全球范围内产生了深远的影响。然而，随着模型规模的指数级增长（参数量从亿级到万亿级），传统 Transformer 中注意力机制的计算复杂度问题日益凸显：
1）二次复杂度挑战：标准注意力机制的 $O(n^2)$计算复杂度导致处理长上下文数据时（如长下文理解、超长文档检索等）面临巨大挑战，导致了高昂的推理成本和较长的回复时间，严重影响了用户体验。
2）GPU 显存限制：注意力矩阵的显存占用随上下文长度线性增长，严重制约模型的可扩展性和应用场景多样性。
3）Test-time scaling 限制：随着思维链数据在大模型广泛应用，思维链数据中文本长度也越来越长，传统 Transformer 在长上下文训练和推理上的效率问题就会被进一步放大，导致推理时间和成本过大。

现有优化方案的局限性：最近涌现出一批新的长上下文高效注意力机制结构用于解决传统 Transformer 中计算复杂度等问题,主要包括稀疏注意力(如 Native sparse attention(NSA)、MoBA）、线性注意力（Mamba2、LightningAttention、RWKV)等。这些方法将传统注意力机制 $O(n^2)$计算复杂度降低至线性或者亚线性，提高了大模型的训练效率，降低了大模型训练和推理的耗时和成本。其中 NSA 和 MoBA 这种稀疏注意力机制，通过分层稀疏化策略，仅关注关键信息，显著降低计算复杂度，但这种稀疏化需要复杂的动态路由策略，工程实现难度较大。而

且稀疏化过程中可能忽略细粒度局部特征，影响对短距离依赖的建模能力。其中线性注意力机制将注意力耗时降低至线性，而且由于避免了显式构建和存储 $O(n^2)$的注意力矩阵，内存消耗大幅减少，使得模型可以在不牺牲性能的前提下处理更长的输入。但由于采用固定 hidden state 来存储历史信息，随着文本长度增加，这种固定 hidden state 降低了大模型对历史信息的检索的准确性，造成了大模型在长文本中检索和捞针能力的下降。因此，设计一种基于长上下文的高效注意力机制的新 Transformer 架构，能够兼顾模型计算效率与性能的帕累托最优显得尤为重要。

目标和产出：

1.论文及专利：发表领域内顶级会议/期刊高质量论文 1-2 篇，申请 1 项专利。
2.技术指标：与现有传统 Transformer 架构相比，性能不下降的情况下，新的注意力机制训练速度提升约 30%，推理速度提升 1 倍。

14. 面向多轮交互的 Agent 化代码生成强化学习框架研究

课题背景：

随着大语言模型(LLM)在代码生成领域取得突破性进展，单轮代码生成已逐步成熟。然而，真实场景的软件开发本质是多轮交互过程：开发者需根据编译错误、测试反馈、需求变更等不断完善、修改代码。因此基于多轮交互的代码开发工具的重要性日益增加。

Agent 化技术通过构建具有规划-执行-反思能力的智能体（Agent），可模拟人类开发者分步拆解任务、自主执行测试、根据错误反馈迭代优化代码。而强化学习（RL） 是实现该闭环的关键——通过构建代码质量、执行成功率、人类偏好等多维奖励函数，驱动 Agent 在交互中持续进化。

近年来，业界已经有诸多代码多轮 Agent 工具证明了在软件开发领域的有效性（如 GitHub Copilot、Codex），蚂蚁自研百灵大语言模型具有较好的单轮代码生成能力，具备结合多轮 Agent 实现真实软件开发的潜力。因此面向多轮交互的 Agent 化代码生成强化学习框架研究具有很高的研究价值和广泛的应用场景。

目标和产出：

1.技术产出:产出高质量多轮 agent 代码强化学习数据 20 万条，并应用在百灵大模型上在 swe-bench、aider 等榜单上取得 20 个百分点的提升。
2.论文及专利：结合百灵大模型的工作，发表领域内顶级会议/期刊高质量论文 1-2 篇，申请 1-2 项专利。

15. 大规模强化学习训练的高效性与稳定性优化算法研究

课题背景：

近年来，大规模强化学习（RL）在复杂决策任务中展现出巨大潜力，特别是在需要长序列推理（Long Chain-of-Thought, CoT）的推理模型。然而，随着模型规模的扩大和任务复杂度的提升，传统强化学习算法在训练效率、稳定性和资源消耗方面面临严峻挑战。

蚂蚁百灵模型作为业界领先的大模型，已在多个榜单验证了其有效性。然而，在长推理链任务中，现有算法仍存在训练不稳定、收敛速度慢、计算资源消耗高等问题。尤其是在分布式训练环境下，如何高效协调多节点计算、减少通信开销、提升数据利用率，同时保证策略的稳定优化，成为亟待解决的关键问题。

因此，本项目旨在研究面向大规模强化学习的高效性与稳定性优化算法，结合蚂蚁百灵大模型的特点，设计可扩展的分布式 RL 训练算法与框架，优化长 CoT 强化学习场景下的训练效率，并降低资源消耗，推动强化学习在复杂任务中的实际落地。

目标和产出：

1.算法创新：提出 1-2 种高效分布式 RL 训练算法，支持长 CoT 任务的稳定优化，训练效率上对比基线 GRPO/DAPO 提升 50%
2.系统实现：基于蚂蚁百灵模型基座实现优化算法，提供开源或内部可用的高性能 RL 训练工具
3.论文及专利：发表领域内顶级会议/期刊高质量论文 1-2 篇，申请 1-2 项专利。
4.应用落地：基于蚂蚁百灵模型基座得到应用

16. 基于扩散大语言模型的复杂交互网页的实时生成技术研究

课题背景：

随着大语言模型（LLM）能力的飞速发展，利用其实时生成前端代码已成为新兴研究热点。然而，当前技术主要局限于实时静态网页开发。当涉及更复杂的场景（如流畅动画、交互式应用乃至轻量级游戏）时，以 GPT 系列、DeepSeek-V3 等为代表的自回归大语言模型（Autoregressive LLM） 面临根本性挑战：
1）速度瓶颈： 逐 Token 生成机制导致高延迟。例如，高性能硬件（如 H200）上，DeepSeek-V3 的典型生成速度仅约 200 tokens/s。这意味着生成一个包含精细状态管理、交互逻辑和动效的复杂交互应用代码（通常需 5000+ tokens）耗时将超过 25 秒，严重破坏了交互流畅性与用户对即时反馈的期待。
2)能力天花板： 现有技术擅长静态或简单交互页面，但对于需要物理引擎模拟、毫秒级响应、复杂状态同步或精细动画控制的高动态内容，自回归模型在结构理解、逻辑连贯性和生成效率上均表现出明显不足，限制了其在沉浸式或游戏化 Web 体验中的应用。
另一方面，新兴的扩散大语言模型（Diffusion LLM） 作为一种并行生成范式，展现出解决实时动态网页生成难题的颠覆性潜力：
3）生成速度快： 扩散模型的并行化生成机制是其核心优势。其迭代去噪过程天然支持高度并行计算，突破自回归模型顺序生成的限制。谷歌的 Gemini Diffusion 理论吞吐量高达 2000 tokens/s，实际演示中仅需约 2.5 秒即可生成含物理引擎的复杂游戏代码（5000 tokens），速度提升达 10 倍，首次实现复杂交互代码的秒级生成（≤10s）。产业实践（如 Mercury Coder 在 H100 上实现 1000+ tokens/s）也验证了其在速度上的显著优势。
4）结构化逻辑鲁棒性增强： 扩散模型学习整个代码序列的联合分布，而非仅依赖前文预测下一 Token。这种全局视角使其在处理需要强逻辑一致性的结构（如嵌套条件分支、事件监听器、状态机）时更具鲁棒性。初步研究（如 Lai et al., 2024）表明，扩散模型在生成含复杂控制流的长代码时，语法正确性和逻辑连贯性错误率更低。

Fast Research 作为一款实时的可视化交互助理，正在推进的实时可视化互动技术研究，是此类技术落地的关键阵地。 因此，深入研究扩散大语言模型（Diffusion LLM）技术，不仅有助于突破复杂动态网页生成的技术瓶颈，也将直接加速 Fast Research 项目的落地和壁垒的构建。

目标和产出：

1.高速的模型：预计能够产出一版速度提升 2 倍及以上的同尺寸模型（预计 32b 左右），并应用到线上。
2.技术产出：AR 模型转 Diffusion 模型的框架，高质量的 Diffusion LLM 的训练数据。
3.论文：发表领域内顶级会议/期刊高质量论文 1-2 篇。

17. 基于多智能体运行时（Runtime）的自我演进算法研究和落地

课题背景：

自 2022 年 OpenAI 发布 ChatGPT 以来，大型语言模型（LLM）引发了全球人工智能领域的重大变革。LLM 在数学、编程、对话系统等通用任务以及游戏、医疗保健、金融等专业领域展现出卓越的能力。近年来，随着模型上下文协议（MCP）等新技术的引入，LLM 与工具的深度融合催生了新一代多智能体系统。AutoGPT、OpenAgent、LangChain、OpenAI Agents SDK 等框架的出现，为复杂任务的自动化解决提供了强有力的支持。

然而，现有的智能体框架在实现自主持续的自我进化方面仍面临诸多挑战。首先，目前主流框架大多采用人工设计的固定架构，尤其是在工具创建和多智能体系统设计方面，缺乏灵活性，难以适应复杂多变的实际需求。其次，这些框架高度依赖于 GPT、Claude 和 Gemini 等闭源 LLM API，而开源模型在复杂推理和多模态真实场景中性能受限。这种对闭源模型的依赖不仅带来了成本和安全风险，也成为制约智能体系统性能提升的关键瓶颈。最为关键的是，现有框架缺乏系统性的方法来帮助智能体发现自身的局限性并通过实践进行学习。根本挑战不仅在于在已知约束条件下解决问题，更在于使智能体能够识别自身的边界并超越这些边界进行演进。

目标和产出：

1.技术指标：通过演进算法迭代，在权威智能体榜单（如，GAIA、BFCL 等），总榜单达 Top3；相同尺寸及以下模型达 Top1；
2.业务赋能：在 Leopard 等实际产品上线智能体功能；以产品 benchmark 为准绳，使用体验与产品其他功能相当。
3.论文及专利：发表领域内顶级会议/期刊高质量论文 1-2 篇（主要聚焦在 ICLR、ICML、NeurIPS 等主流会议），申请 1-2 项专利。

18. 基于大模型稀疏化在智能体应用方向的推理加速技术

课题背景：

RAG，政务出行助手，UI-Agent 等场景，需要大模型服务处理超长文本，多轮对话及大量图片，同时要求大模型服务能够高性能低延时。随着模型规模与序列长度的指数级增长，大规

模矩阵乘法尤其是二次方计算复杂度（$O(n^2)$）的自注意力机制（Self-Attention）成为制约推理效率的核心瓶颈。为达到业务要求，公司现行解决方案往往是堆叠算力，这带来了极大的算力成本，并且有极强的边界效应。因此，研究推理加速技术，改进现有智能体大模型架构和推理流程，是对大模型服务降本增效的根本解决方案。

目前智能体的推理速度受到如下限制：高延迟，低吞吐，高显存：处理长文本，多图片时，点积注意力机制导致推理时延和算力成本激增；长程依赖弱化：智能体模型现有优化技术（如局部窗口注意力，块注意力）难以兼顾全局语义捕捉，导致多跳推理、跨文档分析等任务精度下降。

当前，大模型推理加速的主流技术存在明显局限，比如：静态稀疏方法（如 Longformer 滑动窗口）无法根据输入动态调整计算范围，在复杂语义场景中漏检关键信息；硬件不友好设计：非结构化稀疏模式（如随机注意力）难以利用 GPU 张量核心并行能力，实际加速比不足理论值的 30%；量化与蒸馏虽可压缩模型，但无法根本性解决注意力计算的复杂度问题。

为突破上述瓶颈，亟需一种从结构层面改进智能体模型，硬件协同的稀疏架构，实现：1，计算点积注意力复杂度从 $O(n^2)$降至近 $O(n)$，支持百万级长序列实时推理；2，适配现有主流硬件加速技术。

目标和产出：

1.技术指标：研究面向智能体基于稀疏化技术的基座大模型的推理加速技术。模型推理速度至少提升 10%，且保持模型效果最多下降 1%。
2.算法原型：完成适用于基座大模型的稀疏化的算法设计，交付一套适配于主流软硬件平台的稀疏化的基座大模型的结构及权重文件，包括源代码、权重等。
3.专利：乙方协助甲方和/或其关联公司提交国内、国际专利 1-2 项，交付符合本项目创新点的专利申请号和专利申请相关材料。

（三） AGI Infra

1. 面向新型 LLM 架构的自动分布式并行优化

课题背景：

LLM 训练通常采用 pipeline parallelism (PP)，data parallelism (DP)，tensor parallelism (TP)和 sequence parallelism (SP)中的多种相结合的分布式并行策略以充分利用宝贵的计算资源和加速训练。对于过往的成熟的 LLM 架构，许多手动和自动的优化方法已经能够实现较高的资源利用率。然而，蚂蚁技术研究院在探索创新的 LLM 范式（如扩散语言模型）和 LLM 架构的过程中，由于新的架构带来了计算的不规整，负载均衡的难度增大，导致过往的优化方法失灵，模型训练的资源利用率降低，也无法真正发挥新型 LLM 架构的高效性。

本项目旨在研究针对新型 LLM 范式和架构，自动分析其各模块性能成本特性并生成最优或接近最优分布式并行策略的方法。

目标和产出：

1.针对新型 LLM 范式和架构，自动或半自动地进行分布式并行优化的工具，包括更准确的性能成本建模方法，复杂并行策略空间中的高效搜索，以及可能需要的动态负载均衡等，生成的分布式并行策略 MFU 可以达到或接近 40%，并能够整合进 megatron 等训练框架中。
2.发表领域内顶级会议/期刊高质量论文 1 篇。

2. 大语言模型面向训推一体架构的分布式调度系统和资源优化研究

课题背景：

随着大模型的快速发展，数据中心的智算集群规模也在快速增长。智算集群的资源管理和调度有两个比较大的特征：硬件异构化，以及业务多样化和复杂化。智算集群分布在不同的地域，每个集群会包含不同的加速器硬件，在同一个集群内也会包含不同代际或者型号的加速器；在业务上，智算集群包含预训练、推理、后训练等多种不同的场景，这些场景对应的模型规模、数量、SLO 要求也有很大的不同。

资源调度是站在全局的角度来解决大模型业务负载使用硬件资源的问题，如何在保证业务 SLO 的情况下，充分地利用好集群的资源，提升资源的利用率对于企业来说是非常重要的研究领域。

模型任务的调度需要考虑地域差异、异构硬件分布、模型的资源需求和 SLO、模型在不同硬件的性能指标、模型的并行策略、数据的分布和传输、硬件的资源限制和瓶颈等因素，调度算法的复杂度非常高，是业界和学术界非常重要的研究方向和课题。

目标和产出：

1.论文及专利：结合训推场景的资源优化和能效提升，发表领域内顶级会议/期刊高质量论文 1 篇，申请 1 项专利。
2.技术产出：针对训练推理过程中的资源使用和能效方向，产出落地到蚂蚁内部的调度工程算法，在相同硬件资源的基础上实现任务处理效率 20%的提升。

3. 面向强化学习的大模型推理加速技术

课题背景：

随着大语言模型（LLMs）和强化学习（RL）技术的深度融合，如何高效优化 RL 训练过程中的大模型推理成为关键挑战。RL 任务（如游戏 AI、机器人控制、自动化决策）通常依赖大模型进行状态表示、策略生成或价值估计，但其计算开销大、延迟高、泛化性不足等问题限制了实际应用。本课题面向强化学习中的大模型推理进行软硬协同、深入优化探索。
1）大模型推理需求：面向长序列、大规模模型和数据，结合并行计算、软硬结合优化等，提升模型推理性能，提升吞吐、降低延迟；
2）推理-训练协同优化：探索大模型推理结果如何反馈至 RL 训练过程，探索训推协同的推理优化。

目标和产出：

1.围绕强化学习场景中的大模型推理深入探索，目标产出：1）提出 1-2 种针对 RL 任务的大模型推理优化技术；2）在典型 RL 任务中中实现显著效率提升（如吞吐量↑30%，延迟↓50%）；
2.论文及专利：发表领域内顶级会议或期刊高质量论文至少 1 篇；申请至少 2 项专利。

4. RL 训推一体近似计算优化

课题背景：

随着大模型 CoT reasoning 能力的兴起，后训练 RL 得到了广泛重视。当前 RL 依然面临一些挑战亟需突破性优化，例如：
1）性能挑战：长序列生成（例如~32K 或更长）耗时久，存在长尾，影响效率。
2）迭代速度：训推资源效率较低（特别是稀疏 MoE 模型训练 MFU 较低），影响模型迭代。
3）成本挑战：长序列生成叠加稀疏 MoE 训练门槛较高，需要大量 GPU 资源，成本压力大。

目标和产出：

1.针对 RL 训推场景需求的优化系统实现，在指定典型大模型评测速度和精度，满足以下目标的一项或多项。
1)相比 SOTA，推理吞吐加速>=30%，精度可接受。
2)相比 SOTA，训练吞吐提升>=20%，精度可接受。
3) RL 迭代效率加速 >=25%, 精度可接受。
2.发表领域内顶级会议或期刊高质量论文 1-2 篇。

5. 高性能 Agentic RL 系统研究

课题背景：

蚂蚁通用人工智能（AGI）基础系统 ASystem 当前使用强化学习模式来构建高性能的 AI 持续学习基础系统，希望通过非工作流编排的端到端大模型强化学习训练，支持实时多环境交互生成能力以及大规模、高性能强化学习训练，为大模型在泛化智能体应用场景带来更精准、更高智能水平的多轮交互、多工具编排的能力。

我们希望研究一套具备大规模环境交互（Envs）、多轮交互（Mutil-Turn Acts）、多工具模型调用（Multi-Tools）、部分生成训练（Partial Rollout）的 Agentic RL 系统，能够实现百级工具扩展、万级并发环境交互，并且在规模化性能上远超出主流开源框架，帮助大模型高效训练迭代。

目标和产出：

1.一套高性能 Agentic RL 系统，端到端性能是主流开源框架的 2 倍以上；
2.具备大规模环境交互、多轮交互、多工具模型调用、部分生成训练等能力，实现百级工具扩展、万级并发环境交互；
3.论文及专利：发表领域内顶级会议或期刊高质量论文 1 篇；申请 2 项专利。

6. 面向大规模后训练场景下的训推一体引擎加速

课题背景：

随着大模型训练逐步迈入后训练时代，无论是语言模型、多模态模型，还是智能体的持续训练与优化，其模型复杂度和训练规模均已达到千卡、万卡级别的超大规模水平。
在此背景下，训练成本日益高昂，如何通过训推一体化引擎的深度优化，实现训推全过程的加速，从而有效降低成本开销、提升算法迭代效率，已成为当前算法与工程层面的关键挑战。

希望能引入研究型实习生，共同探索并构建面向大规模后训练场景的高效训推一体引擎，在理论研究与工程实践中推动大模型训练技术的持续突破。

目标和产出：

1.训练 MFU 提高至 25%，逐渐逼近理论上限 28%。
2.训推一体的引擎内只做一次计算，覆盖 generate 和 log_probs 等计算，成本节省 50%。