

计算机前沿技术课程研究报告

李冬晴

2024 年 12 月 30 日

摘要

在协作任务中构建具有自适应行为的代理是多代理系统领域的首要目标。目前开发合作智能体的方法主要依赖于基于学习的方法，其策略泛化在很大程度上取决于他们在训练阶段与之互动的队友的多样性。然而，这种依赖限制了智能体与不熟悉的队友合作时的策略适应能力，这在零样本协调场景中成为一个重大挑战。为了应对这一挑战，作者提出了 ProAgent，这是一种新颖的框架，它利用大型语言模型 (LLM) 来创建能够动态调整其行为以增强与队友合作的主动代理。ProAgent 可以分析当前状态，并从观察中推断队友的意图。然后它会根据队友随后的实际行为更新其信念。此外，ProAgent 表现出高度的模块化和可解释性，使其可以轻松集成到各种协调场景中。在 Overcooked-AI 环境中进行的实验评估揭示了 ProAgent 显著的性能优势，在与 AI 代理合作时优于基于自我博弈和基于群体训练的五种方法。此外，与人类代理模型合作，与当前最先进的方法相比，其性能平均提高了 10% 以上。

关键词：大模型推理；大模型规划；多智能体协调

1 引言

选择该论文复现原因有如下四点：1、创新的框架设计：作者提出了一种名为 ProAgent 的新型框架，旨在通过大型语言模型 (LLMs) 创建能够主动适应并提高团队合作能力的代理。这一框架包括四个核心模块：计划器、验证器、控制器和记忆模块，以及一个信念修正机制。这种高度模块化和可解释性的设计使得 ProAgent 可以轻松集成到各种协调场景中，展示了其在多智能体系统中的巨大潜力。2、优越的性能表现：实验结果表明，ProAgent 在 Overcooked-AI 环境中表现出色，超越了基于自我博弈和基于群体训练的五种方法。此外，与人类代理合作时，其性能平均提高了超过 10%。这些结果证明了 ProAgent 作为一个合作代理在不同场景下的有效性和适应性。3、丰富的研究背景和应用前景：论文详细介绍了大语言模型在长对话、推理和文本生成等领域的应用进展，并探讨了这些模型在个体代理中的潜力。然而，如何将这些模型应用于多智能体协作任务仍是一个未解决的问题。因此，复现这篇论文可以帮助进一步探索这一领域，推动大语言模型在多智能体系统中的创新应用。4、详细的实验设计和评估：论文提供了全面的实验设置和评估方法，包括与不同类型代理的合作能力测试，以及在多种布局下的性能比较。这些实验设计和评估方法为后续研究和开发提供了宝贵的参考，有助于复现和验证该研究的发现。

2 相关工作

一是掌握利用提示语引导大模型有效推理的方法，二是通过更换模型来优化框架性能，三是对论文 [1] 中实现目标的最佳动作函数进行修改。

2.1 提示语引导推理

重点在于探索如何撰写合适的提示语，使大模型能够基于这些提示进行符合逻辑和有效的推理。这涉及到对不同类型问题、不同领域知识的提示语设计和测试，通过大量的实践来找到能让大模型准确输出的提示语规则。

2.2 更换不同模型提升性能

其核心是尝试不同的模型架构，评估每个模型在给定框架下的表现。这需要对多种模型进行训练和测试，比较它们在准确性、效率、泛化能力等方面的差异，进而选择出最能提升框架性能的模型。

2.3 修改最佳动作函数

这里的工作主要是对用于达成目标的最佳动作函数进行调整和优化。需要对函数中的参数、逻辑关系等进行深入分析和修改，通过不断试验来找到能够更高效、更准确地达到目标的动作函数。

3 本文方法

3.1 ProAgent 框架

提出的 ProAgent 框架的概述如图 1 所示，包括协调任务工作流（左）和 ProAgent 管道的内部细节（右）。团队代理的决策回路由外圈蓝色实线箭头表示，ProAgent 的决策过程由中间灰色虚线框和外圈灰色实线箭头表示。ProAgent 通过将初始状态翻译成自然语言开始其操作。然后，Planner 结合存储在 Memory 中的历史信息熟练地分析所提供的语言状态。这个分析过程允许模型辨别队友的意图，并相应地为代理设计高水平的技能。对预测意图的信念将通过信念修正机制进行更新，该机制涉及将其与团队代理随后的实际行为进行比较。对于计划的技能，验证者验证它是否可以在当前状态下执行。在技能失败的情况下，验证者将评估技能的先决条件，并提供遇到问题的详细解释。如果需要，ProAgent 将进入重新规划循环，启动重新校准过程。另一方面，如果该技能被认为是可行的，控制器将进一步将其分解为几个执行低级动作，以便在环境中执行。

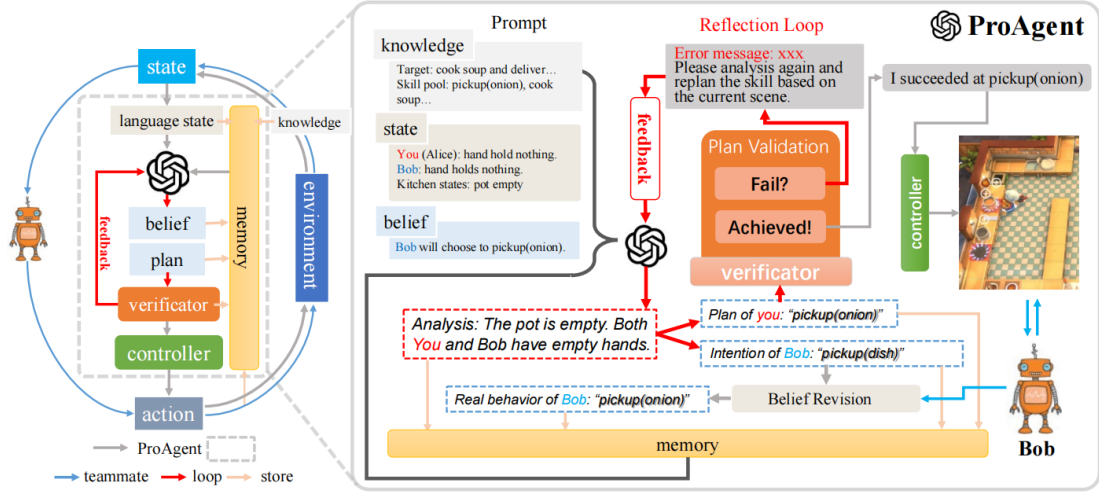


图 1. ProAgent 框架

3.2 关键模块

ProAgent 的流程包括以下几个关键模块：首先，通过知识库和状态关联模块获取与当前任务相关的知识，并将原始张量状态信息转换为基于语言的状态描述，以便大语言模型 (LLM) 能够有效理解。接着，在高级技能规划模块，LLM 规划器基于语言化状态分析当前场景，推测队友代理的意图，并为当前代理规划相应的技能。随后，通过信念修正模块得到进一步校正，以确保对队友代理意图的合理判断。在技能验证与动作执行模块，所选技能由验证模块进行验证，若技能失败则需要重新规划；当技能有效时，控制模块会将其分解为低级别的动作，使 ProAgent 能够高效地与任务或环境交互。最后，在整个过程中，所有与提示、规划、验证和信念修正相关的信息都会存储在记忆模块中，这些存储的知识不仅帮助代理做出明智决策，还能随着时间推移不断调整行为以适应环境变化。

4 复现细节

4.1 与已有开源代码对比

代码开源，复现的时候对比最新的 AI 智能体，比原文更加具有说服力。性能表现方面将 gpt-4o 引入框架并优化到达目标的最佳动作函数中，来提升原本 gpt-3.5 的性能表现。

4.2 实验环境搭建

在 Overcooked - AI 环境中进行实验，以评估 ProAgent 的性能。实验环境搭建过程如下：1、环境选择依据：基于之前关于合作 AI 和人 - AI 合作的研究，选择 Overcooked - AI 作为测试环境，该环境要求两个智能体通过放置食材、煮汤、装汤和送餐等操作合作准备和提供汤品，智能体必须动态分配任务并有效合作。实验使用了该环境中的五个经典布局，包括 Cramped Room、Asymmetric Advantages、Forced Coordination、Coordination Ring 和 Counter Circuit。2、对比算法选择：主要关注基于零样本协调 (ZSC) 方法开发的智能体与不同队友（从不同 AI 智能体到人类）的合作能力。因此，将 ProAgent 与领域中五种常见的替

代算法进行比较，包括自对弈（SP）、基于种群的训练（PBT）、完全合作规划（FCP）、最大熵策略（MEP）和基于合作学习的进化算法（COLE）。3、实验设置与数据收集：将上述六种算法两两组合，构建了 36 对智能体组合（如 SP 与 PBT、PBT 与 SP 等）。对于每对算法组合，进行五次实验（episode），并收集每次实验的回报（episode returns）的平均值和标准差。

4.3 创新点

1、利用大模型的推理和规划技术。2、多智能体协调技术。

5 实验结果分析

在所有五种布局中，优化后的协作 400 步得分都比原文有所提高，且每次提高的分数都是 20。这表明替换大模型且优化到达目标的最佳动作函数后在各个布局下都起到了相同程度的积极作用。实验结果如图 2所示：

协作400步的得分

布局	Cramped Room	Asymmetric Advantages	Coordination Ring	Forced Coordination	Counter Circuit
原文	160	200	120	40	100
优化	180	220	140	60	120

图 2. 实验结果示意

6 总结与展望

在所有五个布局（Cramped Room、Asymmetric Advantages、Coordination Ring、Forced Coordination、Counter Circuit）中，优化后的协作 400 步得分都比原文有所提高，且每次提高的分数均为 20。这表明优化措施在各个布局下都起到了积极作用，能够稳定地提升协作得分。优化效果的一致性较高，说明优化方法可能具有通用性，不受特定布局特性的影响，能够在不同的场景下发挥作用。实验过程中的不足有 1、实验布局局限性。实验仅在五个经典布局中进行，可能无法完全代表所有可能的场景。在实际应用中，可能存在其他未被测试的布局或场景，这些场景下优化方法的效果未知。2、对比算法局限性。仅与五种常见的替代算法进行比较，可能无法涵盖所有可能的算法类型。存在其他未被测试的算法在与优化方法对比时可能表现出不同结果的可能性。3、实验次数有限。对于每对算法组合仅进行五次实验，实

验次数相对较少，可能导致结果存在一定的随机性，无法完全准确地反映优化方法的长期和稳定效果。未来，研究更多不同类型的布局，包括随机生成的布局和实际应用场景中的布局，以验证优化方法在更广泛场景下的有效性。引入更多不同类型的算法进行对比，包括新兴的算法和针对特定场景设计的算法，以全面评估优化方法的竞争力。进行更多次的实验，减少结果的随机性，更准确地评估优化方法的长期和稳定效果。

参考文献

- [1] Ceyao Zhang, Kaijie Yang, Siyi Hu, Zihao Wang, Guanghe Li, Yihang Sun, Cheng Zhang, Zhaowei Zhang, Anji Liu, Song-Chun Zhu, Xiaojun Chang, Junge Zhang, Feng Yin, Yitao Liang, and Yaodong Yang. ProAgent: Building Proactive Cooperative Agents with Large Language Models. *Proc. AAAI Conf. on Artificial Intelligence*, 2024.