

无限记忆，无限探索——开放式环境中的通用智能框架

Kaijie Du
Fudan University
Date: 2024. 4. 24.

1 研究背景与重要性

随着近期LLaMA3模型的开源和大型语言模型领域的竞争加剧，我们见证了人工智能技术在模拟和增强人类语言处理能力方面的显著进步。然而，尽管这些模型在处理和生成自然语言方面展现出强大的能力，它们在持续对话理解和生成方面仍面临记忆容量有限和输出多样性/探索性不足的挑战。当前的大型语言模型，如LLama3，本质上是基于文本的条件反射系统，类似于人类背诵课文，记住成语词汇歇后语。但人类的智能并非仅建立在此基础之上；相反，它是建立在多个智能中枢之间的相互作用和内部言语交互上。这种多智能体的内部交互是思维的真正源泉，也是所谓的“思考”。

我关注了这一方面的很多前沿文献。**我觉得，在全球范围内，CraftJarvis团队依然是最有可能造出AGI的团队（没有之一！）。这是因为CraftJarvis团队所做出的一系列重大突破：**

- **模型的记忆能力：** 现有大型语言模型在持续对话和复杂互动任务中显示出记忆能力有限。但是JARVIS系列装备了一个多模态记忆系统，帮助其使用预训练知识和实际游戏经验来进行计划。
- **多模态输入的整合与处理：** 尽管一些模型已开始尝试处理视觉和语言的输入，但多模态输入的整合和优化处理仍然是一个挑战。JARVIS系列智能体结合了多模态基础模型和语言模型来改善任务理解 and 环境反馈的正确性与一致性。
- **自主学习与自我改进的机制：** 现有大型语言模型大多依赖于大量的预训练，缺乏在实际环境中自我学习和适应的能力。JARVIS则展示了通过自我指导任务提出和执行这一机制来自我改进的能力。这种自我改进的能力是JARVIS能够持续在开放世界游戏中表现出色的关键因素之一。

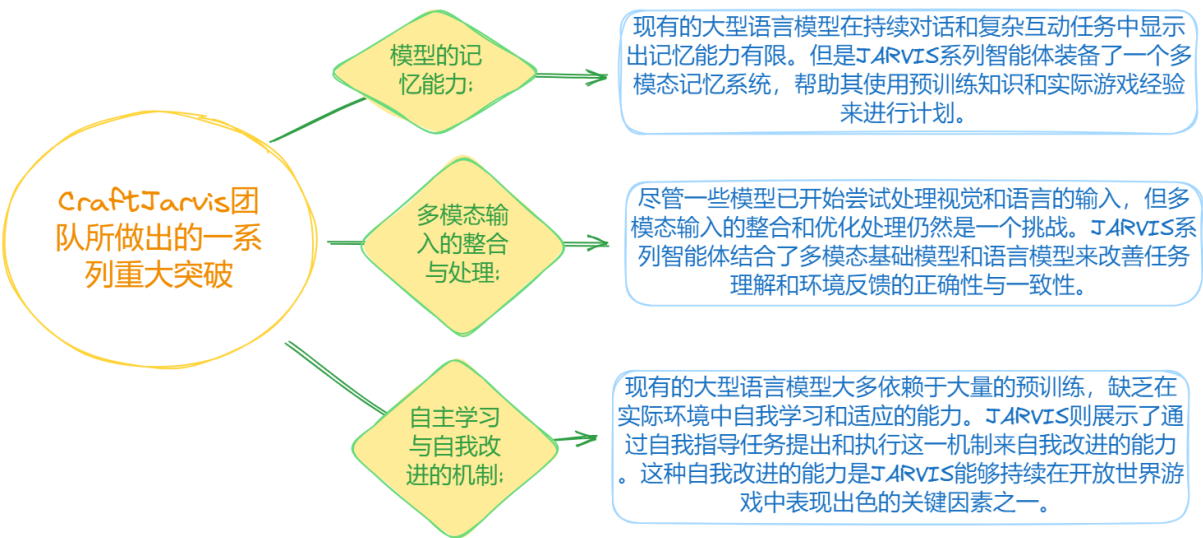


Figure 1: 为什么JARVIS是最接近AGI的智能体

但是JARVIS的记忆处理依然过于直接，庞大的多模态记忆会制约模型的无限制发展。本研究提出通过融合以下人脑记忆功能来突破这些限制，旨在通过构建一个能在复杂互动环境如Minecraft中有效运作的模型，超越传统LLM+多模态记忆+RAG的限制，进一步推动通用人工智能（AGI）的发展。

- **突触可塑性，记忆效用评估：** 人类大脑通过突触可塑性来学习和记忆信息。这是一种神经生物学现象，指的是神经连接的强度根据经验而改变。我们的模型将模拟这一过程，进行记忆效用评估，通过增强频繁使用的神经路径（长时程增强，LTP）和减弱不常用的路径（长时程抑制，LTD）来优化记忆存储。
- **忘记的功能：** 忘记是大脑优化记忆处理的一个重要方面。通过忘记不重要或过时的信息，大脑能够保持记忆存储的经济性并提高处理效率。模型将实现类似的机制，以提高整体性能和适应性。
- **记忆的巩固与重构：** 记忆巩固是将新学的信息从短期记忆转移到长期记忆的过程，通常在睡眠期间发生。记忆重构是指当一个已存储的记忆被重新激活或回忆时，它变得易于修改的过程。我们的模型将模拟这一过程，以支持记忆的长期保存和灵活更新。
- **记忆的抽象和泛化：** 大脑能够从具体经验中抽象出一般规则和模式。模型也将具备这种能力，从具体数据中学习并泛化这些知识，以适应新的情况和挑战。

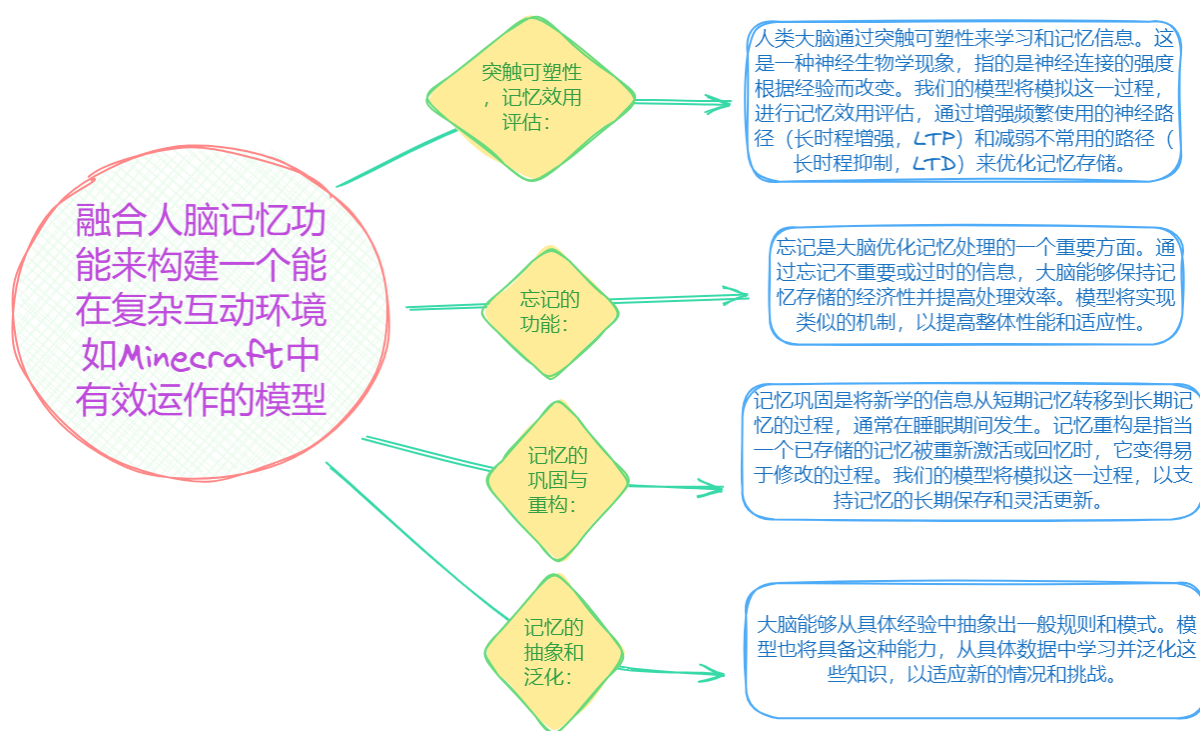


Figure 2: 我将做的工作:智能体的人脑记忆模拟

我认为，目前开源的大型模型已具备支撑AGI诞生的潜力，关键在于如何高效地构建这些模型的交互结构，使得由LLM组成的智能系统转化为真正的AGI。

2 现有的相关工作与文献回顾

2.1 JARVIS-1: Multimodal Agents in Open-World Environments的突破性贡献

JARVIS-1在开放世界多任务智能体领域中代表了一系列突破性的贡献，其创新能力主要包括：

- **多模态输入处理：**JARVIS-1整合并处理来自不同模态的输入，例如视觉数据和文本指令。这增强了其在复杂环境中的适应性和任务执行能力。
- **增强的记忆系统：**通过引入多模态记忆系统，JARVIS-1不仅能存储过往的成功经验，还能在执行新任务时利用这些经验进行更精确的规划和决策。
- **自我改进能力：**通过自我指导生成任务和探索环境，JARVIS-1进行自我学习和改进，显著提高了任务处理的灵活性和效率。
- **执行长期和复杂任务的能力：**与其他智能体相比，JARVIS-1在处理需要长期规划和复杂决策的任务（如在Minecraft中制造钻石镐）上表现出显著的优势。

2.2 GROOT : Learning from Gameplay Videos的突破性贡献

- **新型学习框架：**CraftJarvis团队提出了一个新的学习框架，利用游戏视频作为动态视觉指令，消除了对文本-游戏注释的需求。这种方法大大降低了训练数据准备的复杂性和成本。
- **视频指令编码器：**CraftJarvis团队开发了视频指令编码器，该编码器处理游戏视频以形成结构化的目标空间。这种编码器增强了代理理解和解释复杂任务的能力。
- **编解码器架构：**采用因果变换器基础的编解码器架构实现有效学习和执行动作序列，使模型能够处理《Minecraft》中复杂的任务动态。
- **性能基准测试：**引入“Minecraft技能锻造”基准测试允许对AI在各种任务中的能力进行全面测试，显示模型相对于其他AI代理和人类基准的优越性能。
- **诱导目标空间分析：**通过定性分析，诱导出的目标空间揭示了目标组成和复杂行为合成等新兴属性，突显了模型对游戏策略元素的高级理解。

2.3 VOYAGER: 开放式环境智能体的探索和学习

- **自动课程表：**自动课程表利用GPT-4生成的动态任务推动智能体不断探索和挑战。这种课程设计旨在通过不断提供新任务来激发智能体的好奇心和内在动机，同时确保学习过程中的挑战性和可管理性。
- **技能库：**技能库是VOYAGER的核心组成部分，其中存储了可执行代码的复杂行为。智能体在完成的任务时生成的动作程序被添加到库中，随后可以在面对类似情境时被检索和复用，有效地加速学习过程并增强任务执行的灵活性。
- **迭代提示机制：**迭代提示机制通过环境反馈、执行错误和自我验证的循环反馈来不断优化动作程序。这种方法不仅提高了代码生成的准确性，还通过自我验证确保了任务的成功完成，从而使智能体在没有人类干预的情况下自主学习和适应新任务。

2.4 其他技术：RAG记忆库

2.4.1 llm-agent

在GitHub上的llm-agent项目展示了一个利用向量数据库实现大型语言模型长期记忆的应用。这种方法通过检索增强生成（RAG）技术，使LLM能够从外部数据库检索事实信息，增强其对话能力和信息处理能力。该项目使用mistral-7b-instruct-v0.2.Q4_K_M.gguf模型，并结合了LLAMA.cpp_python接口和chromadb数据库。主要功能包括向数据库添加新信息、查询数据库信息以及实验性的网络搜索功能。GitHub: llm-agent。

2.4.2 MyScale 和LangChain 的集成应用

MyScale 和LangChain 的结合使用为大型语言模型提供了一个强大的结构化存储和检索解决方案。这一系统不仅利用了MyScale 的数据库功能，还整合了LangChain 的SQLChatMessageHistory 组件，有效支持了复杂的数据类型和多租户的需求。此外，MyScale 的向量存储功能与LangChain 的模型管理能力结合，可以高效地进行文本数据的解析和向量化，极大地增强了聊天机器人处理和存储历史数据的能力。这些技术的应用，不仅提升了信息检索的准确性，也为聊天机器人提供了持续的上下文支持，使其在长期对话中更为精准地理解和响应用户需求。

MyScale & LangChain 结合使用案例: <https://myscale.com>

LangChain 的SQLChatMessageHistory 介绍: <https://langchain.com>

3 研究目标

本研究旨在通过集成现有项目的一系列方法和模型，探索和实现在虚拟环境（如Minecraft）中的模拟人类记忆和模拟人类输出的融合。具体研究目标包括：

3.1 模拟人类记忆的实现

- **实现突触可塑性模拟：**开发一个基于神经可塑性原理的学习算法，该算法能够根据交互频率自动调整连接权重。通过集成LangChain和LLAMA3模型，在模型中实施类似于生物神经网络的长时程增强和抑制机制，使得频繁的交互路径增强，不常用的路径减弱，从而优化记忆存储结构和提高信息处理效率。
- **构建忘记机制：**设计一个动态的信息遗忘系统，该系统能够根据信息的时效性和重要性自动调整保留或遗忘数据。这将通过在外数据库中对交互历史进行标记和分类实现，利用机器学习算法来预测信息的保留价值，并定期清理认为不再需要的数据，以模拟人脑中的忘记功能。
- **优化记忆巩固过程：**在模型训练和微调阶段引入类似于人类睡眠过程中的记忆巩固机制。使用定期的“重训练”周期来加强最近和频繁使用的记忆路径，同时利用非监督学习技术在模型的离线阶段重新整理和巩固存储的知识。
- **开发记忆重构与泛化能力：**实现一个记忆重构框架，当模型遇到新信息或被要求回顾旧信息时，可以调整已存储的记忆内容。此外，开发一种从具体实例中抽象出通用规则的机制，使模型能够不仅回忆具体事件，还能在类似情境中应用已学习的知识和技能。

3.2 模拟人类输出的探索

- **扩展模型的输出能力：**利用LangChain的技术栈，尤其是其对不同数据类型和输出格式的支持，扩展LLAMA3模型的能动性。这将使模型能够在更多的场景中（比如除了minecraft以外的简单任务）生成更丰富、更多样化的交互式内容。
- **改进输出指令生成方案：**在JARVIS项目的基础上，增加更多的模拟指令和应用场景，通过增强学习和模式识别算法，自动识别和整合新的用户需求和环境变化，实时更新和扩展模型的输出库。

3.3 虚拟环境中的应用验证

- **全科技树的自动解锁：**开发智能系统以自动完成Minecraft中所有科技树的解锁。这不仅测试了系统的学习和适应能力，还展示了其在处理连续任务中的效率。
- **复杂建筑的自动构建：**利用AI系统自动设计并建造复杂的建筑物，如城堡、高塔等。这将验证系统在空间理解和创造性任务中的应用能力。
- **环境交互的动态适应：**系统需能根据环境变化做出响应，如调整策略以适应突发事件（如生态群系变化、敌对生物出现等）。
- **多任务处理能力：**评估系统在同时处理多个任务时的表现，例如同时进行资源创造（种菜，烧炉子）、探索和防御等活动。
- **用户互动和反馈学习：**系统应能根据用户的指令和反馈进行学习和适应。

4 研究的创新点与潜在影响

本研究的创新之处在于将模拟人类记忆技术结合应用于Minecraft这一虚拟环境中，通过构建一个能够与用户交互，记住智能体与环境/用户之前的所有交互，并不断探索学习的智能系统。此外，研究还将探索这种融合技术是否能真正的实现‘无限性’——无限记忆，无限探索。

5 研究日程表

本研究计划从2024年5月开始，至2025年2月结束。以下是每个阶段的具体任务和目标。

5.1 文献回顾与基础架构设计（2024年5月- 2024年7月）

- 5月：全面回顾相关大型语言模型和记忆模拟的最新研究，熟悉研究框架。利用这段时间与其他研究人员/导师讨论，以验证并敲定我的研究方向和方法。
- 6月：熟悉先前研究中的所有模块，开始开发基于JARVIS的记忆模拟初步框架。
- 7月：开发模拟突触可塑性的算法，设计并实现记忆效用评估与忘记机制。确保设置明确的里程碑和评估标准，这样可以在进入下一阶段前评估进展。

5.2 算法开发与初步实验（2024年8月- 2024年10月）

- 8月：继续研究开发记忆聚类 and 抽象化机制，将复杂的多模态记忆压缩成文本与指令信息。
- 9月：继续研究开发。开始撰写论文初稿，尤其是背景和方法部分。
- 10月：测试模型的泛化能力，全面测试系统性能。。

5.3 系统优化与论文撰写（2024年11月- 2025年2月）

- 11月：根据测试结果优化模型性能。撰写论文后续部分。
- 12月：完成论文定稿并提交至arxiv，将论文最终稿提交至顶会如ICLR/NeurIPS。
- 1月：交流意见，整理研究成果，进行后续研发。
- 2月：开始后续研究内容。(在我的AGI整体研究计划中，我还列举一些后续的AGI实现关键的初步想法)

DETAILED PLAN		May 2024	June 2024	July 2024	August 2024	September 2024	October 2024	November 2024	December 2024	January 2025	February 2025
Literature Review and Infrastructure Design	Conduct a comprehensive review of recent studies including JARVIS, mineCLIP	✓									
	Get acquainted with all modules	✓	✓								
	Develop algorithms to simulate synaptic plasticity.		✓	✓	✓	✓					
Algorithm Development and Preliminary Experiments	Continue developing memory clustering and abstraction mechanisms			✓	✓	✓					
	drafting the paper				✓	✓	✓	✓			
	Test the model's generalization ability					✓	✓	✓			
System Optimization	Optimize the model performance					✓	✓	✓			
Paper Writing	Complete the final draft of the paper and submit it to arxiv								✓		
	submit the final draft to top conferences such as ICLR/NeurIPS								✓		
discussing and debating revisions with reviewers	respond to reviewer comments.									✓	✓
	engage in discussions with reviewers									✓	✓

Figure 3: Detailed Plan

6 参考文献

- [1]Wang, Z., Cai, S., Liu, A., Jin, Y., Hou, J., Zhang, B., ... & Liang, Y. (2023). Jarvis-1: Open-world multi-task agents with memory-augmented multimodal language models. arXiv preprint arXiv:2311.05997
- [2]GitHub: llm-agent <https://github.com/AkiRusProd/llm-agent> .
- [3]Cai, S., Zhang, B., Wang, Z., Ma, X., Liu, A., & Liang, Y. (2023). Groot: Learning to follow instructions by watching gameplay videos. arXiv preprint arXiv:2310.08235.
- [4]Wang, G., Xie, Y., Jiang, Y., Mandlekar, A., Xiao, C., Zhu, Y., ... & Anandkumar, A. (2023). Voyager: An open-ended embodied agent with large language models. arXiv preprint arXiv:2305.16291.
- [5]Zhang, C., Yang, K., Hu, S., Wang, Z., Li, G., Sun, Y., ...& Yang, Y. (2024, March). ProAgent: building proactive cooperative agents with large language models. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 38, No. 16, pp. 17591-17599).
- [6]Lin, H., Wang, Z., Ma, J., & Liang, Y. (2023). Mcu: A task-centric framework for open-ended agent evaluation in minecraft. arXiv preprint arXiv:2310.08367.