



2023下半年：

- **更便宜、更快的GPT-4**：这是OpenAI的首要任务。总体而言，OpenAI的目标是尽可能地降低“智能成本”，因此他们将努力继续降低API的成本；
- **更长的上下文字数限制**：未来高达100万个token的上下文字数限制是可预期的；
- **微调API**：将扩展到最新模型，但具体形式会根据开发人员实际需求进行调整；
- **有状态的API**：目前调用聊天API时，必须反复提交相同的对话历史记录并再次支付相同数量的token成本，未来则会有一个能记得对话历史的迭代；

2024年：

- **多模态**：这是GPT-4发布时演示的一部分，但在更多GPU资源上线之前，暂时无法让所有用户体验到；

迄今为止规模最大的图灵测试实验

Probability of Correct Guess	
Overall	68%
When Partner is a Bot	60%
When Partner is Human	73%

Table 1: Probability of correct guess by partner type.

AI 21实验室完成了迄今为止规模最大的图灵测试实验，150多万参与者进行了超过1000万次对话，其中73%的时间里能够正确识别人类，但在与机器人交谈时只有60%的时间识别正确。

论文：<https://arxiv.org/pdf/2305.20010.pdf>

项目：<https://www.ai21.com/blog/human-or-not-results>

OpenAI 用「过程监督」大幅提升GPT-4数学推理能力消除「幻觉」

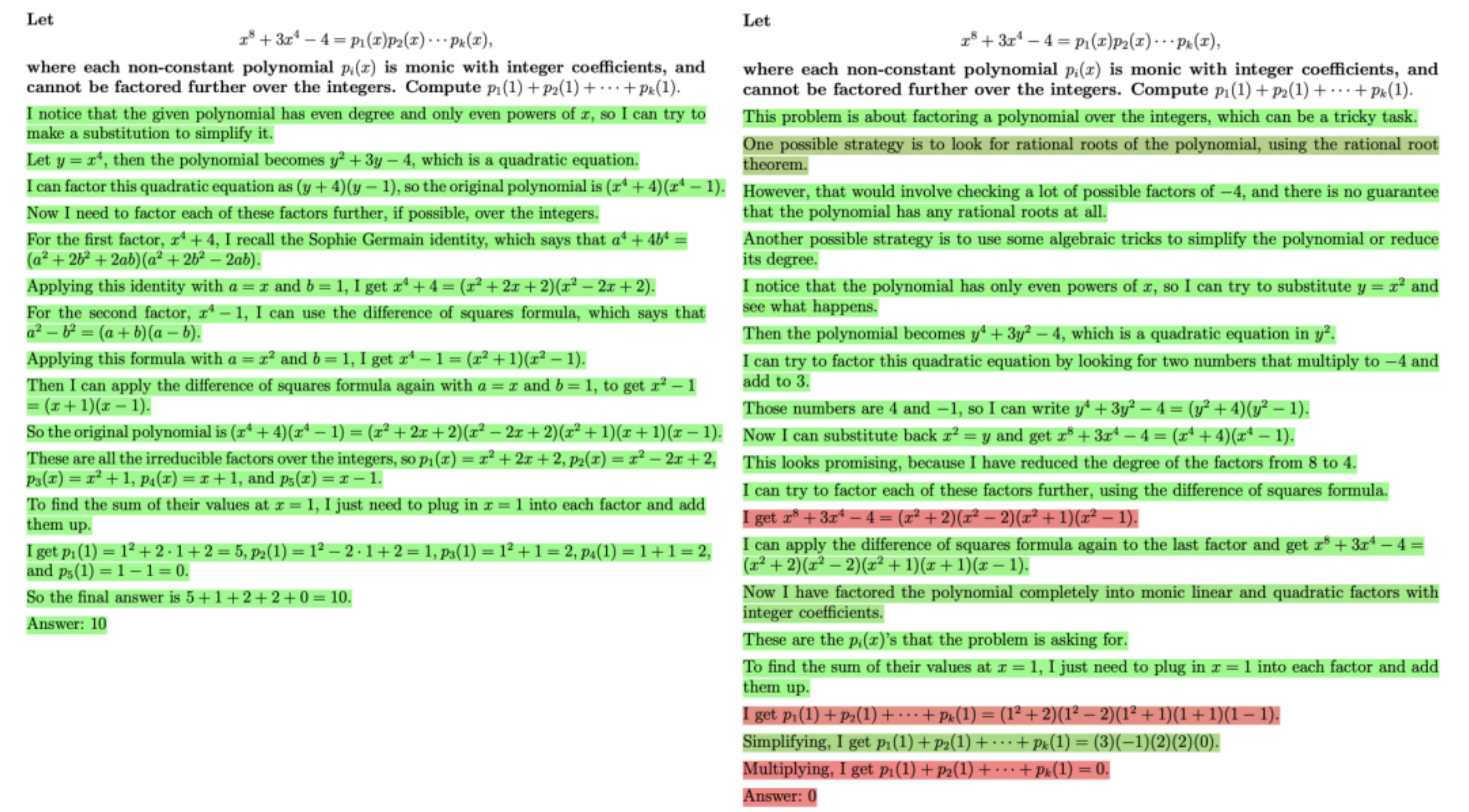


Figure 2: Two solutions to the same problem, graded by the PRM. The solution on the left is correct while the solution on the right is incorrect. A green background indicates a high PRM score, and a red background indicates a low score. The PRM correctly identifies the mistake in the incorrect solution.

论文引入了一种被称为“过程监督”（PRM）的新方法，以提高人工智能模型中的数学推理能力。该技术侧重于推理过程，训练模型逐步解决问题，并为其解决方案提供解释。它比以前的工作有了显著的改进，扩展了大型语言模型处理更复杂数学问题的能力

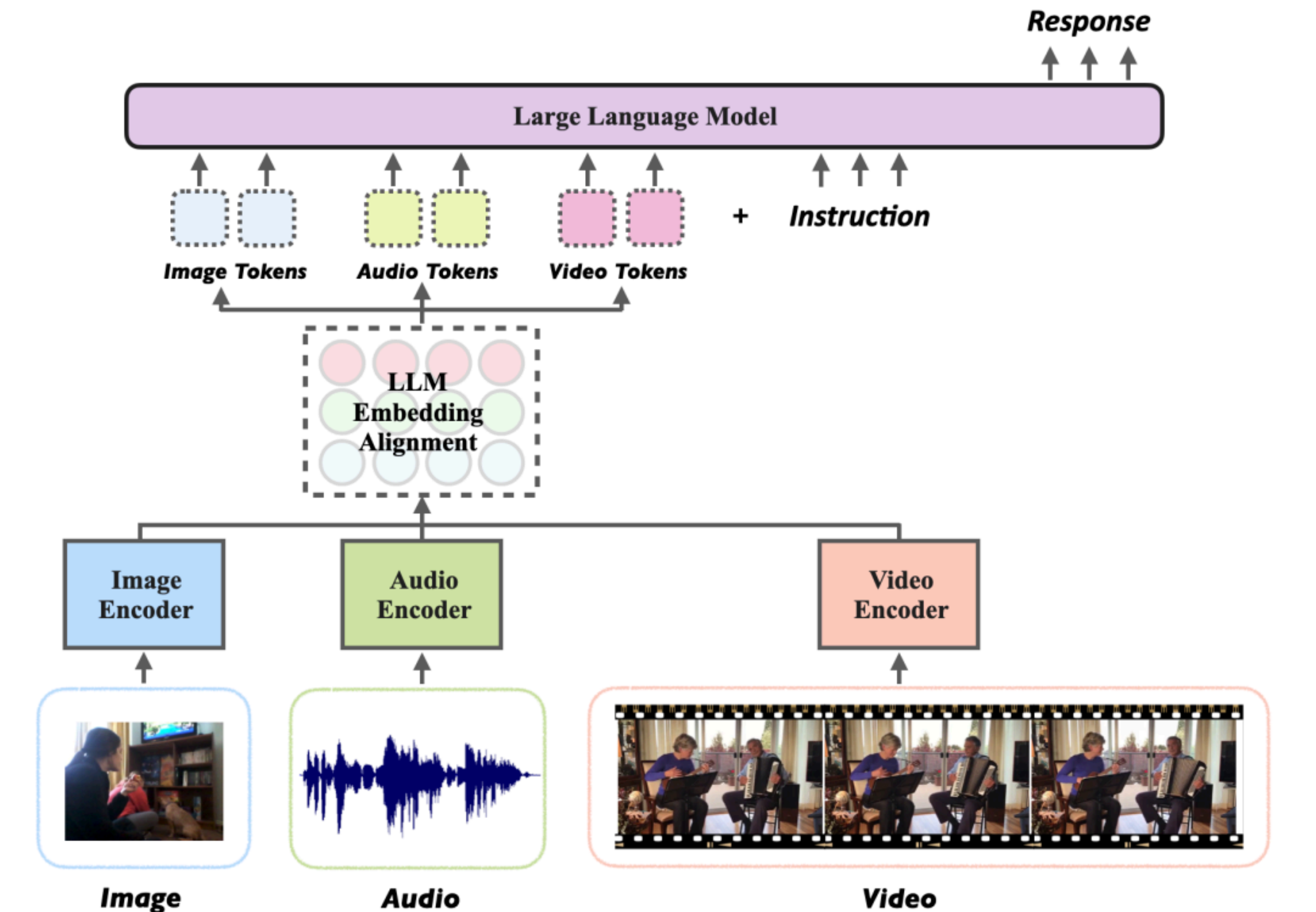
论文：https://cdn.openai.com/improving-mathematical-reasoning-with-process-supervision/Lets_Verify_Step_by_Step.pdf

LangChain课程

本课程面向刚接触LangChain的开发人员，LangChain是一个强大的开源框架，用于快速搭建基于LLM的应用程序。作者曾在12周内利用LangChain搭建了12个应用，并且每个月新增百万用户。

课程：<https://github.com/SamurAIGPT/langchain-course>

Macaw-LLM



Macaw-LLM是一个探索性项目，通过无缝结合图像、视频、音频和文本数据来开创多模态式语言建模，并基于CLIP、Whisper和LLaMA等基础模型构建。