

一种基于上下文学习的迭代式机器人任务规划方法

王浩岑 郑朝文 赵卓峰

北方工业大学 信息学院 大规模流数据重点实验室, 北京, 100144;

摘要: 大型语言模型 (LLMs) 通过分析大量文本数据, 积累了丰富的语言使用模式和世界知识, 成为理解和生成自然语言的关键工具。在机器人领域, 如何将这些知识转化为机器人的任务规划能力, 是一个研究热点。可以利用预训练的语言模型, 通过适当的提示, 将抽象任务分解为机器人可执行的动作。且在无额外训练的情况下, LLMs 可以生成可执行的行动计划, 如提示词足够充分, LLMs 能够将高级任务转化为逻辑连贯的行动步骤。以上, 不难看出 LLMs 在机器人行动规划中的潜力。然而, 这些行动计划在现实世界中的可执行性仍面临挑战, 意味着将语言模型知识转化为机器人行动需要进一步研究。本文提出一种基于上下文学习、利用大模型的能力, 进一步探索和改进前人方法, 提高机器人任务规划的准确性和可执行性的方法。

关键词: 上下文学习; 大模型; 任务规划

DOI: 10. 69979/3041-0673. 25. 08. 021

绪论

随着人工智能技术的快速发展, 机器人任务规划已成为智能机器人领域的关键技术之一。任务规划的目标是根据抽象的指令和当前环境状态, 规划出一系列可执行的子任务, 以实现复杂目标。在非结构化环境中的应用, 如家庭服务、工业生产、安防救援等领域, 任务规划问题显得尤为重要。传统的任务规划方法, 如基于经典规划模型和马尔可夫决策过程 (MDP) 模型, 存在效率和泛化能力的局限性。近年来, 大语言模型 (LLM) 的集成改变了机器人领域, 使机器人能够更熟练地进行交流、理解和推理, 尤其是在上下文学习 (In-Context Learning, ICL) 方面展现出了强大的新学习范式。

尽管上下文学习为机器人任务规划带来了新的可能性, 但目前仍存在一些挑战和问题。首先, 传统的基于 LLM 的任务规划方法存在效率问题, 如高代价的令牌消耗和冗余的错误修正, 这些问题限制了其在大规模测试和应用中的可扩展性。其次, 现有的上下文学习方法在面对环境变化时泛化能力较弱, 尤其是在动态、非确定性问题中。此外, 如何将上下文学习与机器人的实际物理操作相结合, 以及如何处理大规模问题中的高维状态空间, 也是当前研究中需要解决的问题。最后, 上下文学习在机器人任务规划中的可解释性和安全性也是研究中需要关注的重要方面。

本文针对大模型生成的任务规划准确率不高的问题, 提出了一种基于上下文学习的迭代式机器人任务规划方法。该方法的设计思想是通过设计提示词来提高大模型的能力, 首先使用动态样本选择算法得到合适的样本示例并输入大模型, 将得到的机器人行动指令转换,

并利用大模型反思得到准确且可执行性高的生成行动序列。这一方法融合了自然语言处理、动作规划和上下文学习技术, 为机器人任务规划领域提供了新的视角和解决方案。

1 相关工作

目前研究的机器人利用大模型的任务规划方法主要有以下三种, 下面分别进行介绍。

1.1 零样本计划者的语言模型

在论文“Language models as zero-shot planners: Extracting actionable knowledge for embodied agents”中^[1], 作者提出了一种方法, 通过利用大型语言模型 (LLMs) 的内部知识, 将高级自然语言任务描述转化为可执行的中级计划。这种方法通过精心设计的提示, 使得语言模型能够在没有额外训练的情况下, 生成任务的行动步骤。为了提高这些计划的可执行性, 研究者们进一步开发了一个程序, 该程序基于现有示范, 将中级计划语义上转化为具体的可执行动作。在 Virtual Home 环境的评估中^[2], 这种方法显示出了较高的可执行性, 并通过人类评估揭示了在可执行性和正确性之间取得的平衡, 从而证明了从语言模型中提取可执行知识的有效性。

1.2 基于机器人可供性语言的执行方法

SayCan^[3]这篇论文中提出的机器人任务规划方法是一种结合了大型语言模型 (LLM) 和机器人技能的框架, 用于将高级指令转化为机器人可以执行的具体动作。该方法首先通过语言模型评估每个技能对完成高级指令的有用性, 同时考虑技能执行成功的概率, 从而选择最

合适的技能来执行。这个过程涉及到一个对话式的规划结构，用户给出高级指令，语言模型生成明确的操作序列。SayCan 通过迭代选择任务并将其添加到指令中，直到任务完成。此外，SayCan 还能够处理多步骤任务，并且能够根据环境反馈动态调整计划，以适应错误恢复和任务的连续性。这种方法在真实世界的厨房环境中得到了验证，展现出良好的泛化能力和执行准确性。

1.3 其他机器人任务规划方法

在机器人任务规划领域，研究者们提出了多种方法来提升规划的效率和适应性^[4]。首先，一种方法是通过自然语言交互来指导机器人行为，利用语言模型来理解自然语言请求并驱动机器人学习物理世界的知识，这种方法通过语义分析理解人类命令，从而提高机器人的指令执行能力。其次，另一种策略是采用层次化的方法，将复杂任务分解为更简单的子任务，通过 AI 任务规划器生成候选计划框架，然后由动作规划器验证其几何可行性，这种方法通过分层规划和动作生成，有效处理具有复杂几何约束和组合复杂性的任务。最后，还有研究提出了基于约束的 TAMP 方法^[5]，将任务排序和低级几何约束结合起来，通过特殊求解器处理约束关系，这种方法通过构建约束网络，有效处理运动规划中的几何约束问题^[6]。

2 基于上下文学习的迭代式规划方法

2.1 方法概述

为了提高大模型在机器人任务规划中的准确度及

可执行性，本文提出一种基于上下文学习的迭代式规划方法（如图 1），其核心思想是通过大语言模型（LLM）的上下文学习能力，结合提示词工程、大模型反思与链式迭代生成等方法，逐步生成可执行的任务规划方案。方法分为以下三个阶段：

（1）选择初始提示词：将 Sample 向量化，与 Query 计算语义距离，得到最佳提示词，按照提示词框架构建出提示词 A1。

（2）第一步生成与大模型自我反思：将提示词 A1 输入到大语言模型中，得到待选 Step1 序列（Step1, Step1', Step1''）；将该序列返回大语言模型，使用新的提示词使大模型挑选最佳候选步骤决策，得到最佳 Step1；对其使用 SBERT 模型，翻译获得可执行的 Step1。

（3）任务序列生成与错误判定：采用递归思想，将上一步中得到的可执行的 Step1 与提示词 A1 组合，得到新的提示词 A2，重新将其输入进大语言模型，得到 Step2，同样使用 SBERT 模型对其进行翻译，得到可执行的 Step2，对 Step1、Step2 序列进行错误判定，如有明显错误则返回第二阶段，否则继续执行；将上一步中得到的可执行的 Step2 与提示词 A2 组合，得到新的提示词 A3，重新将其输入进大语言模型，得到 Step3，同样使用 SBERT 模型对其进行翻译，得到可执行的 Step3，对 Step1、Step2、Step3 序列进行错误判定，如有明显错误则返回第二阶段，否则继续执行……按此方法执行 k 次，得到任务序列（Step1, Step2, Step3, …, Stepk），即为最终目标任务序列。

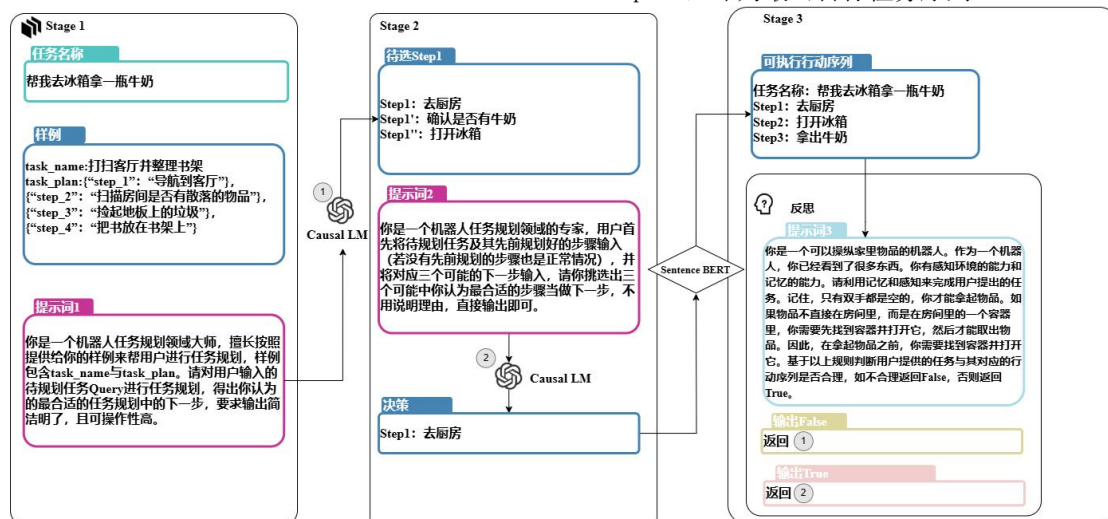


图 1 方法图

使用该方法得到的一串原子动作即为任务规划结果。

2.2 初始提示词选取

本研究采用动态示例选择算法（Dynamic Sample S

election Algorithm）来实现提示任务的优化选取。该算法的核心机制是通过预训练的翻译语言模型（Translation LM）计算查询任务（query task）与候选任务集

之间的语义相似度矩阵,并基于最大相似度准则选取最优示例,并选出最佳 Step1。具体而言,算法首先构建任务-行动步骤的映射关系,随后在嵌入空间中计算查询任务与各候选任务的曼哈顿距离(Manhattan Distance)^[7],选取相似度最高的任务-行动步骤对作为提示学习的示范样本(demonstration sample),将其作为提示词提交给 Causal LM 后生成待选 Step1 序列,最终在待选序列中选出最优 Step1。这一过程既确保了示例选择与目标任务在语义空间上的高度一致性,从而提升了提示学习的有效性,又能高效的选出最佳 Step1。

2.3 第一步生成与大模型自我反思

步骤生成提示词用来生成动作序列的第一步(Step 1)。

设计目标:通过定义标准化的任务描述模板,生成行动序列的第一步。

提示词设计为“你是一个机器人任务规划领域大师,擅长按照提供给你的样例来帮用户进行任务规划,样例包含 task_name 与 task_plan。请对用户输入的待规划任务 Query 进行任务规划,得出你认为的最合适的任务规划中的下一步,要求输出简洁明了,且可操作性高。”

用户输入包含样例 Sample 和待规划任务 Query,其中 Sample 包含 task_name 与 task_plan,而 Query 中可能包含之前步骤中规划好的 Step1, Step2, ..., Stepj。

在上一步中得到了三个待选第一步,为了挑选最佳 Step1,需通过设计候选步骤决策提示词让大模型挑选得到结果。

设计目标:通过定义标准化的任务描述模板,让大模型决策出最佳 Step1。

提示词设计为“你是一个机器人任务规划领域的专家,用户将待规划任务及其对应三个可能的初始步骤输入,请你挑选出三个可能中你认为最合适的初始步骤,不用说明理由,直接输出即可。”

用户输入包含 Query 和 First Step,其中 First Step 包含三个待选 Step1,而 Query 包含待规划任务和先前规划好的步骤序列 Step1...Stepj。

2.4 任务序列生成与错误判定

在之前的机器人任务规划相关研究领域中,往往忽视了环境信息对操作的可执行性的影响。如“从冰箱中拿取牛奶”任务,机器人可能会因没有执行“打开冰箱门”操作,导致任务失败。原因是没有环境因素的介入,机器人难以判断当前状态是否符合任务需求。

为了判断当前行动序列是否符合当前环境信息和状态,避免浪费资源,设计基于环境状态的动作序列反

思对已有行动序列做具体查询,结合机器人当前的环境对生成的行动序列进行评估与反思,并重新规划冗余和错误的行动步骤。

反思内容包括:1. 环境变化的反思:物品位置变化:检测任务执行后物品的位置是否发生变化。2. 机器人自身限制的反思:-机器人只有两只手,因此累计最多能拿两个物品-机器人与物品的距离,太远拿不到物品 3. 常识性反思:-家具的门,在打开后无法继续打开,在关闭后无法继续关闭-在封闭容器内的物品无法直接拿到,需要将容器先打开。

提示词设计为:你是一个可以操纵家里物品的机器人。作为一个机器人,你已经看到了很多东西。你有感知环境的能力和记忆的能力。请利用记忆和感知来完成用户提出的任务。记住,只有双手都是空的,你才能拿起物品。如果物品不直接在房间里,而是在房间里的一个容器里,你需要先找到容器并打开它,然后才能取出物品。因此,在拿起物品之前,你需要找到容器并打开它。基于以上规则判断用户提供的任务与其对应的行动序列是否合理,如不合理返回 False,否则返回 True。

3 实验

3.1 实验设置

3.1.1 实验环境

VirtualHome 作为一个多智能体模拟平台,模拟家庭活动并支持复杂交互,如操作电器和移动物体。它具备交互式物体、人形代理、多模态数据输出等功能,并允许程序化生成和物理模拟,适用于活动理解、AI 代理训练和人机交互研究。VirtualHome 的技术核心在于其程序化活动表示,通过 Unity 和 Python 实现的模拟器,以及 Python API,便于用户交互和 AI 代理训练。这个平台为研究提供了一个逼真的测试环境,以优化基于上下文学习的机器人任务规划方法。

具体而言,研究试图揭示,关于如何执行复杂任务(例如“做早餐”),是否能够转化为一系列具体行动,这些行动可以在环境中执行(例如“打开冰箱”,“拿牛奶”,“关闭冰箱”等)。为此,本文采用了最近提出的 VirtualHome 环境。该环境因其能够支持丰富多样的互动,模拟类似于自然人类活动的开放式任务,而成为研究的理想选择。VirtualHome 环境模拟了家庭环境中的大量真实人类活动,并通过定义为动词-宾语结构的具身动作来支持这些活动的执行^[8]。

VirtualHome 中室内环境共有 7 种,在每种环境中都有 4 个房间:厨房、卧室、浴室、客厅。

3.1.2 实验数据集

首先,让所有的 action 与 arg 相连接,得到 121*

32=3872 个原始任务。其次,对这 3872 个原始任务进行人工筛选,去掉明显不符合逻辑的原始任务(如 turn on sofa, put on water 等),剩余 1407 个初始任务。

使用示例样本采样算法,对原有数据集进行改造。从 1407 个初始任务中,挑选出最具代表性的 200 个精选任务,平分生成测试集和训练集,分别将其整理为 JSON 格式文件以便实验使用。

3.2 实验结果

实验结果令人鼓舞。使用上下文学习的迭代式规划

方法,我们的系统在准确性方面实现了显著提升,达到了 72.45% 的准确率。这一结果不仅超越了先前实验中的任务规划准确度,也证明了该方法在理解和生成复杂任务行动计划方面的潜力。此外,系统的可执行性也得到了改善,达到了 14.29%。尽管这一结果仍有提升空间,但它标志着向实现高度自动化的机器人行动规划迈出了坚实的一步。

对任务进行分类后,使用 ChatGlm-6B 模型得到的对不同任务得到的准确率和可执率如表 1。

表 1 不同任务表现表

任务类型	put_dishwasher	put_fridge	prepare_snack	setup_table	prepare_food
准确度	74.52%	65.20%	78.81%	68.35%	67.1%
可执行性	16.3%	10.64%	3.25%	21.7%	11.33%

可以看出,对不同类型的任务,模型得到的准确度和其可执行性有所不同。

消融实验:将动态选择方法去掉后,得到的 ChatGlm-6B 模型准确率为 43.5%,可执行性为 20.1%。

4 研究结论

实验结果证实了上下文学习的迭代式规划方法在机器人行动规划领域的应用价值,同时也为我们未来的研究工作提供了新的方向。我们相信,通过进一步的优化和调整,我们的方法有望在未来的机器人任务规划领域中发挥更加关键的作用。

尽管实验结果中的执行率有所下降,不过这一小小的代价是值得的,因为它换来了准确率的大幅提升。这一结果不仅证明了该方法在处理复杂任务时的潜力,也为未来的研究工作提供了新的方向。

5 不足与未来展望

目前实验仍停留在虚拟环境中,对于真实世界中的机器人,是否能够通过使用上下文学习的迭代式规划方法来提高任务规划的准确率及可执行性仍然未知,待后续工作继续研究。

参考文献

[1] Huang W, Abbeel P, Pathak D, et al. Language models as zero-shot planners: Extracting actionable knowledge for embodied agents[C]//International conference on machine learning. PMLR, 2022: 9118-9147.

[2] Puig X, Ra K, Boben M, et al. Virtualhome: Simulating household activities via programs [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 84

94-8502.

[3] Ahn M, Brohan A, Brown N, et al. Do as i can, not as i say: Grounding language in robotic affordances[J]. arxiv preprint arxiv:2204.01691, 2022.

[4] Liu R, Wei J, Gu S S, et al. Mind's eye: Grounded language model reasoning through simulation[J]. arxiv preprint arxiv:2210.05359, 2022.

[5] Song C H, Wu J, Washington C, et al. Llm-planner: Few-shot grounded planning for embodied agents with large language models[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2023: 2998-3009.

[6] Gu Y, Deng X, Su Y. Don't Generate, Discriminate: A Proposal for Grounding Language Models to Real-World Environments[J]. arxiv preprint arxiv:2212.09736, 2022.

[7] Nottingham K, Ammanabrolu P, Suhr A, et al. Do embodied agents dream of pixelated sheep: Embodied decision making using language guided world modelling[C]//International Conference on Machine Learning. PMLR, 2023: 26311-26325.

[8] Wang Z, Cai S, Chen G, et al. Describe, explain, plan and select: Interactive planning with large language models enables open-world multi-task agents[J]. arxiv preprint arxiv:2302.01560, 2023.

作者简介:王浩岑(1999.12-)男,满族,辽宁省锦州市人,北方工业大学,硕士研究生在读,研究方向:人工智能。