【神经符号 AI 讲座】提升机器人智能的实践——基于大型语言模型的任务规划增强研究

编者按：上海交通大学神经符号人工智能暑期学校开学典礼暨社区研讨会，于 2024 年 6 月 24-25 日在上海交通大学召开。上海交通大学张翌盛在会上作了题为《**提升机器人智能的实践——基于大型语言模型的任务规划增强研究**》的社区共享项目研究进展报告。任务规划能力对于智能机器人在现实世界中的自主操作至关重要。然而，传统的基于规划领域定义语言（PDDL）的方法往往面临组合爆炸和规划时间不理想的问题。为此，我们提出了一种创新的方法，通过大型语言模型（LLM）指导 PDDL 规划器的搜索过程，从而实现 LLM 增强机器人的任务规划能力。具体而言，LLM 通过学习的启发式方法引导 PDDL 规划器的搜索过程，并提供约束推理以有效减少搜索空间。为应对 LLM 的潜在缺陷，我们在执行阶段引入了一个验证机制，以确保计划的正确性。我们通过真实的报废汽车电池拆卸场景对所提出的方法进行了评估，实验结果表明，将 LLM 整合到规划流程中能够显著提升规划效率和可扩展性，同时保持规划的有效性。这项研究为将语言模型与经典方法相结合，以提升机器人在实际应用中的智能水平提供了一个极具前景的方向。

一、大语言模型领域的突破性进展与研究范式的革新

在当今数字化和 AI 化的时代，先进的人工智能与机器人技术的融合不仅是一种趋势，更是一种必然。我们的工作重点是研究一种新颖的机器人任务规划方法，通过利用大语言模型的强大功能，显著提高其任务执行的效率和效果。

尤其是 2022 年以来，大语言模型领域的革命性进展令人瞩目。许多领域都尝试将这些模型应用于解决一般性任务，并展现出模型规模扩大所带来的巨大优势。然而，在处理涉及算术或符号推理的问题时，直接将问题作为输入往往无法达到理想的效果。研究表明，通过使用预期的输入输出方式进行提示，能够使模型进行上下文类的小样本学习，从而提升模型在各类任务上的表现。在大量实践中，我们发现，当要求大语言模型进行分布式推理时，它们往往能够产生更稳健的解决方案，尤其是在处理上述任务规划问题时。

在此基础上，一些研究提出了“思维链”（Chain of Thought, COT）这样的提

示词方法。通过在输入输出范式之间插入中间的推理步骤，引导模型进行分布推理。而“思维树”（Tree of Thought, TOT）方法则是在 COT 的基础上进行了改进，生成树状结构的思维步骤，并引入了一个自我评估机制。通过小样本学习对中间推理步骤进行量化评估，随后通过搜索算法获得最可靠的思路链。

大语言模型应用于机器人领域

目前，已有一些研究尝试将大语言模型应用于机器人感知、控制以及任务和运动规划中，如图 1 所示。例如，微软的研究人员通过定义一组高级的机器人 API 或函数库，成功利用 ChatGPT 自编程的方式执行机器人的操纵任务。其他方法通过开发用于机器人操纵的定制大语言模型来实现任务和运动规划。以 VIMA 为例，它以多模态提示为输入，使具有隐含智能的机器人代理能够理解指令并执行相应的任务。这些模型似乎具有很好的泛化能力，但它们仍然无法克服大语言模型幻觉，故障概率也无法满足实际需求，特别是在工业场景中。

ChatGPT for Robotics[1] (Microsoft)　　VIMA: General Robot Manipulation with Multimodal Prompts[2] (ICML 2023)

[1]S. Vemprala, R. Bonatti, A. Bucker, and A. Kapoor, "Chatgpt for robotics: Design principles and model abilities," Microsoft, Tech. Rep. MSR-TR-2023-8, February 2023.
[2]Y. Jiang, A. Gupta, Z. Zhang, G. Wang, Y. Dou, Y. Chen, L. Fei-Fei, A. Anandkumar, Y. Zhu, and L. Fan, "Vima: General robot manipulation with multimodal prompts," 2023.

图 1.大语言模型应用于机器人感知、控制以及任务和运动规划中

我们的研究主要聚焦于**电动汽车的电池拆卸任务**。随着电动汽车行业的快速发展，提高电动汽车拆卸效率势在必行。然而，由于废旧电池的设计、品牌、完整性以及报废时间等变量众多，拆卸过程往往比装配过程更加复杂。目前，自动拆卸尚不可行，人们开始采用人机混合的方式来提高拆卸效率。具体来说，机器人承担诸如拆卸紧固件之类的简单重复性任务，而需要高灵活性的任务则由人类承担。然而，在长期使用的汽车动力电池上执行拆解任务，对于机器人而言，依然是一项充满挑战性的工作。

二、基于神经符号 AI 的机器人任务和运动规划框架

与自动装配流水线中的机器人不同，拆解流水线中的机器人无法完全通过预编程的动作完成任务，需要具备自主性、可解释性和稳健性。在我们之前的工作

中，提出了基于神经符号 AI 的任务和运动规划架构，以解决非结构化条件下的人机混合拆卸流程中的不确定性问题。

如图 2 所示，我们采用的神经符号 AI 框架形成了一个闭环的机器人具身智能系统，该系统通过融合符号系统的推理能力和神经系统的感知与学习能力，将推理和决策与感知和控制紧密结合。

- 为了全面表征系统状态，我们创新性地引入了**神经谓词**。一个神经谓词就是一个神经网络。多模态感知信息作为这个神经网络的输入，将传感器采集的信号映射至准符号化的状态空间，从而实现了对系统状态的精准描述。
- 为了执行紧固件拆卸任务，我们预先设计了一组受工人手工拆卸操作启发的**动作原语**。这些动作原语由 PDDL 进行定义，并根据相关参数、前提条件和预期的执行效果进行了描述。

机器人在执行任务时，首先会感知环境。在规划系统中，神经谓词集合被表示为状态节点，然后根据系统状态通过广度优先搜索等逻辑搜索算法，得到需要按顺序执行的动作原语序列。为了保证任务执行的安全性和可靠性，在每次机器人执行任务前，系统会不断检查当前状态。如果感知到的当前状态与系统预期不一致，系统将重新规划。

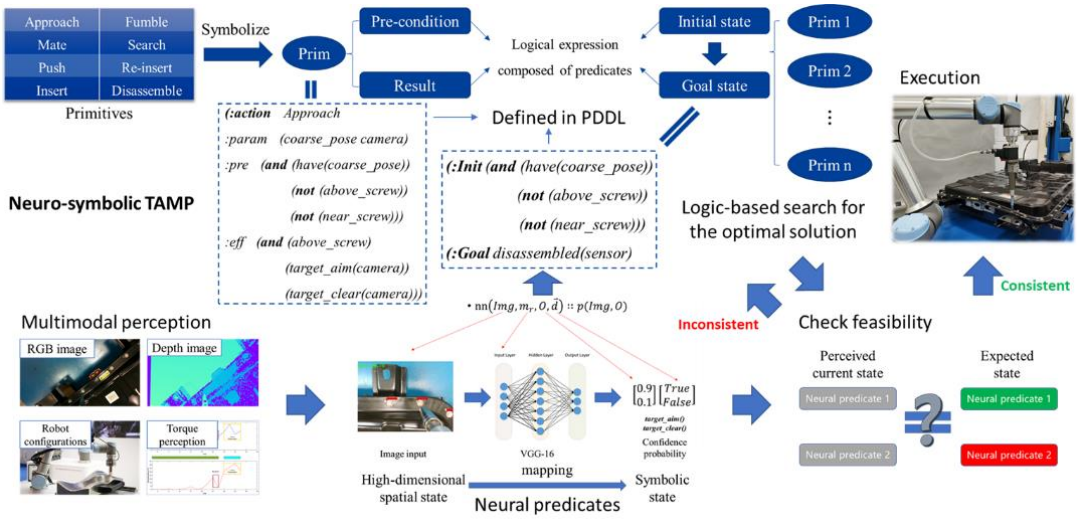

图 2.神经符号 AI 框架形成了一个闭环的机器人具身智能系统

在我们之前的工作中，采用了基于先入先出思想的广度优先搜索算法，生成操作数最少的解决方案。然而，当谓词和原语的数量增加时，这种算法的复杂度会大幅增加，不利于频繁地实施推理。我们的理想是让机器人通过对外部环境的

感知完全自主地进行规划和执行，因此，自然而然，我们探索了一些替代现有任务规划模块的选项。

三、大语言模型增强的任务规划框架

最近，将流行的大语言模型应用于机器人的任务规划引起了我们的兴趣。通过大语言模型可以让机器人自行做出决策，而不是通过人类预设计的算法。因此，我们开发了一种由大语言模型（如 GPT 系列）增强的任务规划方法，其架构如下图 3 所示，它包括一个推理引擎和一个与工作场景实时交互的验证引擎。对于推理引擎，我们将在后面详细介绍。而验证引擎则是我们考虑到现实中的不确定性，特别是人工智能方法的概率性和场景中的边缘案例而引入的。人工智能方法在处理不确定性和随机性问题时，采用概率论和统计学的方法来建模和推理。这种方法允许系统在面对不完全信息或噪声数据时，通过概率分布来表示和处理不确定性，从而做出更为合理和可靠的决策。AI 系统通常是基于大量数据进行训练的，那些在设计或测试过程中不常见、极端或特殊的边缘案例可能会导致系统出现意外的行为或错误。因此，引入了这样一个与工作场景实时交互的验证引擎，提高系统的适应性和稳健性。由于我们的系统源于 PDDL 范式，验证引擎在执行前会对每个原语的预定义前提条件进行二次检查，以确保只有在所有条件都满足时，才会执行相应的动作原语，否则将重新规划。

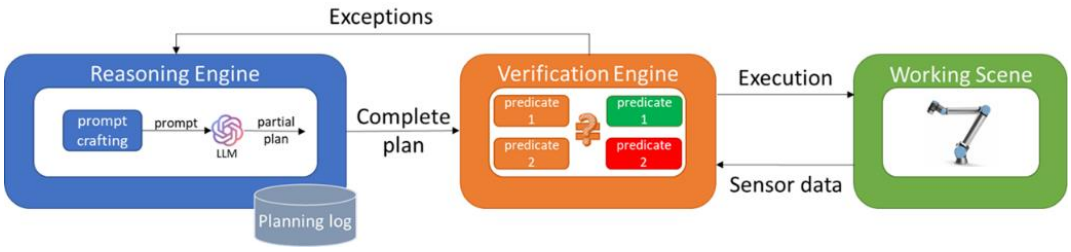

图 3.由大语言模型增强的任务规划架构

由于动态工作环境带来的感知能力等因素的不确定性，紧固件的拆解任务可以看作是一个由许多原语组成的非固定过程的长程任务（Long Horizon Task）。在我们之前的工作中，可信的感知是通过神经谓词实现的。在此基础上，任务规划需要通过可靠有效的提示来进行指导。

许多研究表明，大语言模型在复杂的零样本推理任务上表现不佳，即使是简单的少样本学习，也可能不足以满足应用层面的需求。受到思维树（TOT）的启

发，我们提出了一个由三层框架构建的推理引擎，如图 4 所示。

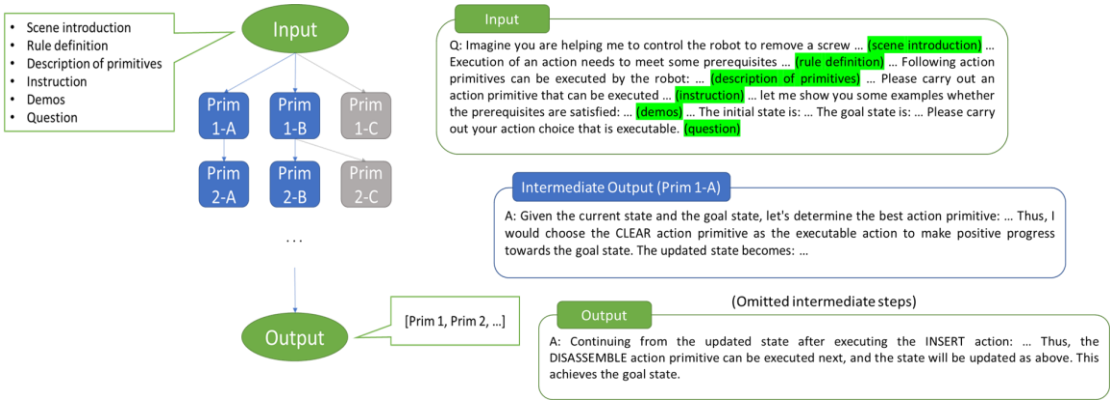

图 4.由大语言模型增强的任务规划推理引擎框架。左边是推理引擎的结构，由三层框架构建的思维树；右边是一个简单的单步规划示例，解释说明推理引擎的结构。

图 4 左边所示是**推理引擎的结构**。与现有的 COT 等采样完整的思维链方法不同，我们生成一颗保持一定大小的思维树。其中每个思维节点由一个连贯的原语序列组成，作为解决问题的中间阶段表示。我们使用精心设计的**提示词**来生成行动计划，其中包括对**神经谓词和动作原语的描述、对工作场景的介绍、对任务规划规则的定义、任务执行的指示**以及推理引擎中包含的**正样本和负样本样例**。其中，**任务规划规则的定义**和**任务执行的指示**，是用于直观地向语言模型传达 PDDL 的概念、任务目标和相应的实现手段。而给出**正样本和负样本样例**的目的，是提供必要的少样本学习事例。我们发现，神经谓词和动作原语是推理引擎的关键要素，因此，究竟是使用**自然语言描述**还是采用更契合 PDDL 范式的**键-值对**形式更为有效是一个值得探究的问题。在后续的实验中，我们对这两者都进行了测试，并且也对少样本学习的效果进行了消融研究。

图 4 右边所示是一个简单的单步规划示例，解释说明推理引擎的结构。Input 阶段给定提示词，我们生成了一个由语言组成的思维节点。然后，不断迭代思维节点的生成，直至生成一个新的任务，直至达到所需的目标状态。

图 5 所示是大语言模型增强的任务规划推理引擎结构。它是按照三层提示词自主生成和评估解决方案的机制设计的。在这个引擎中，我们通过大语言模型的提示生成思维树，并进行检查和评估。这种结构允许语言模型通过基于小样本学习取得的决策评估分数，自主决定是否保留或丢弃当前生成的思维节点。它通过持续的单步推理实现类人的思维方式，最终生成深思熟虑的方案以达到解决问题

的目的。在可行性检查器中，如果提案被判断为不可行，它将直接舍弃；如果被判断为可行，则予以保留。实际上，在 PDDL 规划中，如果它被定义为一个合法方案，我们就会予以保留。

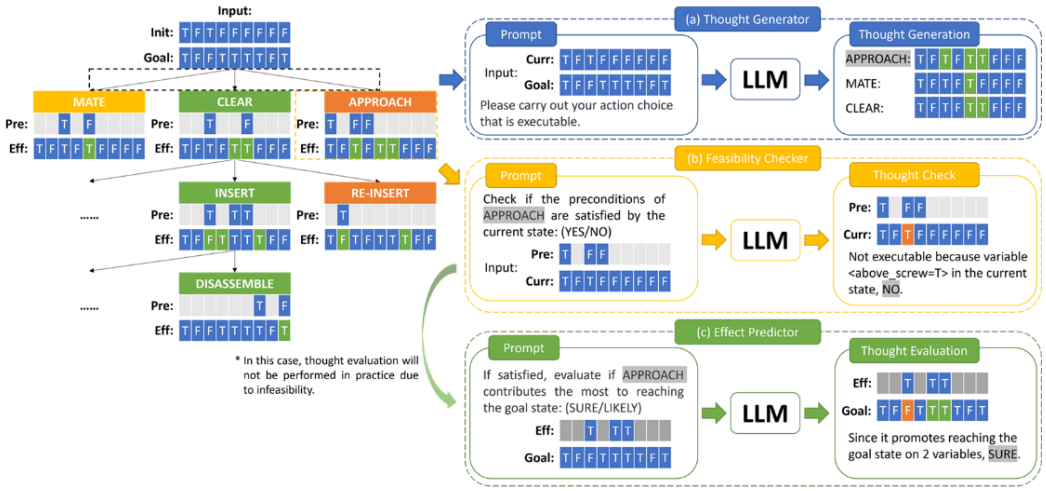

图 5.大语言模型增强的任务规划推理引擎结构

在效果预测器评估中，我们进一步评估该提案对接近目标状态的贡献，或者说它的有效性。如果被判断为对接近目标状态具有重大贡献，我们将给予更高的评估分数；如果它只改变了一个对应的神经谓词或状态，我们则认为它可能促进接近目标状态，并可能得到一个相对较低的评估分数。

这个框架的另一个好处是，它可以插入各种现有的逻辑搜索算法，如广度优先搜索和深度优先搜索算法。例如，当任务搜索空间非常大时，我们可以使用深度优先搜索快速得到一个可执行的实例；当我们需要节省机器人的任务执行时间时，我们可以使用广度优先搜索生成一条操作数最短的操作路径，从而节省整体执行时间。此外，凭借其对自我生成提案的评估机制所带来的稳健性，我们还可以通过直接命令 LLM 一次性自主生成具有最短步骤的完整规划。这样，我们就可以通过直接访问大语言模型，而不是通过预先编程或学习的方法来实现启发式搜索，从而使系统更接近**自主敏捷、可信具身**的智能化目标。

四、实验结果与讨论

为了测试推理引擎的性能，我们结合神经符号任务与运动规划架构，在 GPT-3.5 和 GPT-4 上进行了一系列仿真和真机实验。

PDDL 描述方式的有效性实验

为了探索如何让大语言模型更好地理解 PDDL 概念，我们选择了**自然语言**描述和更加简洁的**键-值对**描述两种方案来表达神经谓词，并进一步描述整个系统状态。

Primitive	Pre-condition	Result	Input Sensor	PDDL Definition	Function Description
Approach	1. there is a coarse pose of the screw to be disassembled acquired via visual perception; 2. the sleeve is not near above the screw; 3. the sleeve is not close to the screw.	1. the sleeve is near above the screw; 2. the sleeve is aligned with the screw; 3. there are no obstacles near the screw.	RGB-D camera	**(:action** Approach :param (coarse_pose camera) :pre **(and** (have(coarse_pose)) **(not** (above_screw)) **(not** (near_screw))) :eff **(and** (above_screw) (target_aim(camera)) (target_clear(camera)))	This primitive moves the robot arm's nut runner so that it's close above the target screw for the subsequent actions.

图 6. Approach（靠近）动作原语的自然语言描述和 PDDL 键-值对描述

给定系统的当前状态和目标状态，我们在思维生成器中进行单步采样过程。在零样本的情况下（即，我们不提供任何操作实例的正负样本）将思维节点是否可行作为评判标准。我们从八个合法的系统状态中随机选择一个作为系统的初始状态，然后在每种描述（自然语言描述和键-值对描述）下分别采样 400 组样本。对于每个特定的模型（GPT-3.5 或 GPT-4），我们发现使用自然语言描述生成的规划可行性略高于**键-值对**描述，但差异并不显著。因此，为了实验的便利性，我们直接使用了**键-值**对形式。

提示词组件的消融实验

对于依赖于大语言模型的逻辑推理方法，提示词的设计至关重要。在我们的规划提示词中，主要包括六个部分，其中，原语描述、指令和问题部分是不可或缺的。为了分析其余组件的作用，我们在 GPT-3.5 上进行了消融研究。在每种情况下，我们执行了 160 个随机样本的单步规划，每个样本包含 3 个合法示例，结果如图 7 所示。

Prompt Component			Success Rate
Scene introduction	Rule definition	Exemplars	
-	-	-	30.6%
✓	-	-	38.1%
-	✓	-	23.1%
-	-	✓	56.9%
✓	✓	-	24.4%
✓	-	✓	50.6%
-	✓	✓	51.9%
✓	✓	✓	58.1%

图 7.提示词组件的消融实验结果。正负样本的引入带来了性能提升。

结果表明，实验场景和任务规划规则等叙述性内容的加入并没有显著促进可行解决方案的生成。相反，正负样本的引入带来了高达 58%的性能提升。也就是说，大语言模型的性能提升最关键的要素在于提供的样本数量。然而，当样本数量增加到一定程度时，简单地增加示例数量并不能进一步提高规划的成功率。因此，我们提出的包括检查和评估机制的推理引擎架构是必要的。

可行性检查器和效果预测器的测试

在评估推理引擎的整体性能之前，我们对可行性检查器和效果预测器分别进行了独立的模拟测试。由于它们实际上充当了生成思维方案的分类器，我们在实践中为每个类别分别提供了一个示例作为参考。我们随机生成了 720 个思维节点，可行性检查器的准确率高达 99%。对于效果预测器，我们将动作原语的效果对接近目标状态的贡献作为评估准则。我们发现，预测器的准确率约为 64%。然而，它并不会影响生成规划的可行性，只会影响执行方案的效率。

真实场景中的机器人实验

在进行仿真实验后，我们将推理引擎与神经符号任务与运动规划架构连接起来，在真实场景中进行实验，以测试其在 GPT-3.5 中的综合性能。我们使用了类似广度优先搜索的搜索策略，但限制了搜索空间。在搜索树的每一层，我们得到了两个不同的可行解时，停止思维节点的生成。如果不满足条件，我们最多生成五个思维节点。实践中证明，这足以确保至少一个可行解决方案。在逐层的迭代过程中，最多保持两个可行的思维节点。最终，我们将最短的思维链路转化为任务规划的解决方案，输出到机械臂执行。

通过与工作场景交互的验证引擎，我们保证了动作原语可以稳健执行。也就是说，我们通过一系列传感器判断当前神经谓词的状态，在最终执行前作为最后一道“关卡”，保证了原语的可行性。推理引擎在仅给定一个样例的情况下，仍然取得了非常优异的性能表现，如图 8 所示。在 160 组实验中，其成功率达到了 96%，远高于无监督的单步规划和基于大语言模型的思维链提示方法。在实践中，通过使用验证引擎，我们系统的最终任务执行成功率达到了 100%。

五、结束语

我们通过无缝集成大语言模型和逻辑推理，提供了一个可靠且高效的机器人任务与运动规划系统。我们可以通过调整每次迭代中的步骤数量，适应不同的大语言模型，从而满足部署需求。验证引擎的存在使我们不必担心大模型的“幻觉”问题。然而，我们目前仅将其用于相对简单的拆卸螺钉场景。由于神经谓词和原语的数量有限，我们还不能进行非常准确的效率比较。未来我们将在更复杂的场景中进行广泛研究。

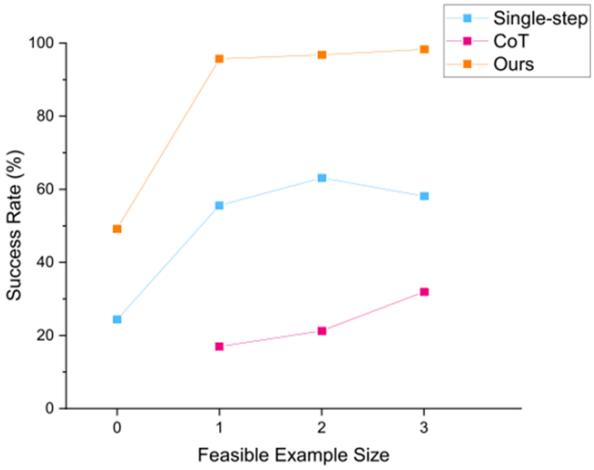

图 8.不同方法真机实验的综合性能比较

我们计划引入更多类型和不同锈蚀状态的紧固件，涉及条件和类型识别，以及快换装置的交互。我们还希望让机器人参与更多的汽车动力电池组件的拆卸工作，如电池模组的抓取、母线排的拆卸等。我们还将引入更多的神经谓词和原语，并通过输入任务描述，引导大语言模型生成新的原语，或对现有原语做出调整，这些都是我们未来工作的努力方向。

（责任编辑 曹晓舟，审核 刘永光）