大模型工具学习进展与挑战

林衍凯
中国人民大学高瓴人工智能学院, 北京，100872，中国
yankailin@ruc.edu.cn

摘要

本论文综述了大模型工具学习的最新进展与挑战。工具作为人类智慧和能力的延伸，在提升生产力和解决问题方面至关重要。随着大语言模型（Large Language Models）的突破，工具学习得到了广泛关注，通过动态调用外部工具，显著增强了模型解决复杂问题的能力。

本文介绍了一个通用的大模型工具学习框架，包括控制器、工具集、环境和感知器四个核心组件。我们详细探讨了四个关键问题：意图理解、规划、工具使用和记忆管理。在意图理解方面，模型需要准确解析用户的输入和隐含意图。规划能力使模型能够将复杂任务分解为可执行的子任务。工具使用方面，介绍了示范学习、教程学习和探索学习三种主要训练策略，通过观察人类示范、阅读工具手册和直接探索来提升模型能力。记忆管理方面，提出了动态记忆管理和信息优先级管理等方法，以提高模型处理复杂任务的效率和准确性。

本文分析了当前大模型工具学习的研究进展和每个领域的挑战，为未来研究提供了有价值的见解。希望通过这篇综述，能帮助研究人员和开发者更好地理解和推进大模型工具学习领域的发展。

关键词： 工具学习；预训练模型

Challenges and Advances in Tool Learning with Foundation Models

Yankai Lin
Gaoling School of Artificial Intelligence, Renmin University of China, Beijing 100872, China
yankailin@ruc.edu.cn

Abstract

This paper reviews the latest developments and challenges in the field of tool learning with foundation models. Tools, as extensions of human intelligence and capabilities, play a crucial role in enhancing productivity and problem-solving abilities. With the breakthroughs in Large Language Models (LLMs), tool learning has gained widespread attention, significantly enhancing the model's ability to solve complex problems through dynamic interaction with external tools.

We propose a general framework for tool learning with foundation models, consisting of four core components: the controller, tool set, environment, and perceiver. This paper explores four key issues: intent understanding, planning, tool use, and memory management. For intent understanding, models need to accurately interpret user

第二十三届中国计算语言学大会论文集，第34页-第47页，太原，中国，2024年7月25日至28日。
卷2：前沿综述

inputs and underlying intentions. Planning capabilities enable models to break down complex tasks into executable sub-tasks. In terms of tool use, we introduce three main training strategies: demonstration learning, tutorial learning, and exploratory learning, which enhance model capabilities through observing human demonstrations, reading tool manuals, and direct exploration, respectively. For memory management, we propose methods such as dynamic memory management and information prioritization to improve the model's efficiency and accuracy in handling complex tasks.

This paper analyzes the current research progress in tool learning with foundation models, highlighting the challenges in each area and evaluating the effectiveness and limitations of existing solutions. By providing valuable insights for future research and applications, we aim to help researchers and developers better understand and advance the field of tool learning with foundation models.

Keywords: Tool Learning , Pre-trained Models

1 引言

工具是人类智慧与能力的延伸，旨在提升生产力、效率和解决问题的能力。从文明诞生之初，工具便成为我们生活中不可或缺的一部分 (Washburn, 1960)。工具的发明和使用，源于人类克服自身局限、探索未知领域的渴望。随着技术进步，我们能够完成越来越复杂的任务，释放出更多的时间和资源去追求更加宏伟的目标。工具不仅是文化和社会实践的重要基石，还极大地改变了我们的学习、交流、工作和娱乐方式，使其变得更加便捷和互动 (Gibson et al., 1993)。人类在工具发明和使用中发挥了关键作用，这无疑是智慧的鲜明体现 (Shumaker et al., 2011)。随着人工智能（AI）的快速发展，一个重要的问题是：人工智能是否具备与人类智能同样的工具学习能力?

掌握工具操作的前提是对工具功能的深刻理解，以及理解用户意图、进行规划和推理的能力。在预训练模型出现之前，进行以工具为导向的人工智能研究面临巨大挑战。尽管某些基础工具可以使用浅层统计模型或深度神经模型来实现调用 (Pomerleau, 1988; Mnih et al., 2013; Akkaya et al., 2019)，但它们的调用效果和稳定性仍不足以满足实际应用的需求，更不用说在各种工具间的泛化能力。这主要是由于传统监督学习在捕捉工具使用中的复杂操作方面存在局限性，而强化学习等试错类范式在掌握工具使用所需的庞大决策空间时也显得力不从心。总的来说，早期的人工智能研究在工具学习方面的根本限制在于模型能力的不足。

今天，我们正处于一个新的技术复兴期，由大语言模型（Large Language Models）的突破所驱动。大语言模型如GPT-4 (Achiam et al., 2023)展示了在自然语言处理（NLP）任务中的卓越能力 (El-Kassas et al., 2021; Zhang et al., 2024; Yang et al., 2018; Kwiatkowski et al., 2019)。然而，尽管这些模型能力惊人，它们在处理复杂计算和提供准确、及时的信息时仍存在挑战，因其依赖固定的参数知识，这常常导致"幻觉"现象，即生成看似合理但事实上错误或过时的回答 (Mallen et al., 2022; Vu et al., 2023; Ji et al., 2023; Zhang et al., 2023; **?**)。

随着大模型能力的不断增强，工具学习 (Qin et al., 2023a)范式被提出，期望大模型能够像人类一样熟练使用工具来解决复杂问题。工具学习通过允许模型动态调用外部工具，不仅提升了大模型的解决问题能力，还拓宽了其功能范围。例如，大语言模型可以使用计算器进行复杂计算以增强其数值计算能力，也通过天气API获取实时天气更新 (Pan et al., 2023; Wang et al., 2024)。调用工具可以显著提高模型响应用户查询的准确性，促进了更有效和可靠的用户互动。随着这一领域的不断发展，大模型工具学习有望在未来的人工智能领域中发挥关键作用，提供更灵活和适应性的解决方案 (Parisi et al., 2022; Karpas et al., 2022; Nakano et al., 2021; Surís et al., 2023)。

在过去的一年里，随着大模型的崛起，工具学习的研究也迅速增加。在实际应用中，GPT-4通过调用插件解决其知识限制，增强其能力，并将插件返回的结果与其内部知识结合，为用户生成更好的响应 (Achiam et al., 2023)。在研究领域，许多研究集中在评估大模型的工具学习能力以及如何增强这种能力 (Qin et al., 2023b; Xu et al., 2023; Gao et al., 2024; Zhao et al., 2024)。鉴于工具学习在大模型领域中的日益关注和快速发展，本文旨在回顾大模型工具学习最新的进展和挑战，以帮助研究人员和产业开发者了解当前的进展。

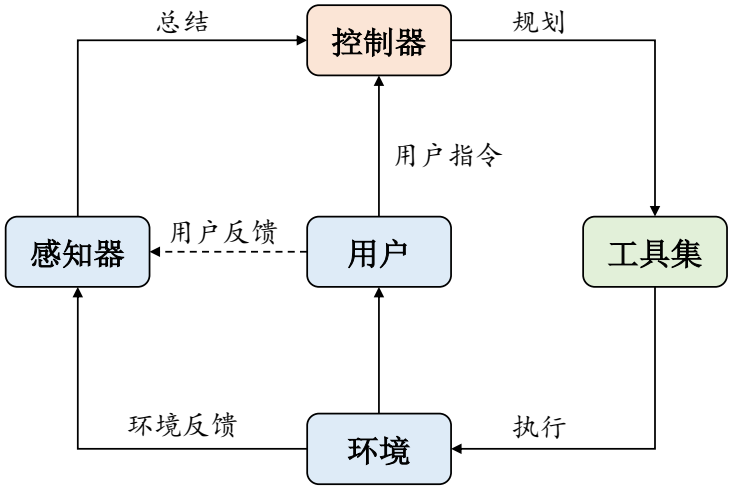

Figure 1: 大模型工具学习框架示意图。

本文首先基于现有工作总结出一个通用的大模型工具学习框架，包括控制器（通常使用大模型建模）、工具集、环境、感知器和人类。在统一框架的基础上，我们以回顾现有的大模型工具学习工作突出其核心研究问题。整个大模型工具学习过程从用户指令开始，模型需要为工具执行制定可执行的计划。为了将用户指令与适当的工具连接起来，模型首先需要学习理解指令背后的用户意图，在此基础上，将复杂任务分解为若干子任务，并有效地用适当的工具完成每个子任务。本文将以典型工作为例介绍其现有研究进展与挑战。

2 大模型工具学习框架

为了更全面地理解工具学习的核心挑战和未来方向，我们在 (Qin et al., 2023a)一文中提出了大模型工具学习的通用框架。具体而言，如图 1所示，我们将工具学习框架分为四个组成部分：包括控制器、工具集、感知器和环境。我们首先介绍工具学习过程中的每个组成部分：

- 控制器（Controller）控制器$\mathcal{C}$ 是大模型工具学习框架的核心，通常基于大模型来实现。控制器的任务是理解用户意图，并制定一个可行且精确的计划来使用工具满足用户需求。控制器需要理解用户的意图以及这些意图与可用工具之间的关系，然后制定一个选择适当工具的计划。对于复杂任务，控制器可能需要将其分解为多个子任务，这需要大模型具备强大的规划和推理能力。

- **工具集（Tool Set）**:工具集$\mathcal{T} = \mathcal{T}_1, \mathcal{T}_2, \cdots$ 是大模型工具学习的基础，包含一系列具有不同功能的工具。每个工具都有不同的接口。我们主要以应用程序编程接口（API）为例，说明如何与工具进行交互。API可以被定义为任何能够将基础模型的输出作为输入的功能。

- **环境（Environment）**：环境$\mathcal{E}$ 是工具执行的平台，为工具执行提供必要的基础设施，并向感知器提供工具执行结果。环境可以是虚拟的，也可以是真实的。虚拟环境具有易于访问和复制的优势，适合成本效益高的模型训练。然而，虚拟环境可能无法完全反映现实世界的复杂性，导致过拟合和泛化问题 (Hansen et al., 2021)。相比之下，真实环境提供了更具现实感的场景，但访问难度和成本较高。

- **感知器（Perceiver）**：感知器$\mathcal{P}$ 负责处理来自用户和环境的反馈，并生成信息的摘要供控制器使用。简单的反馈处理包括将用户和环境的反馈进行拼接或使用预定义模板进行格式化。总结后的反馈将传递给控制器，以辅助其决策。通过这些反馈，控制器可以判断生成的计划是否有效，以及执行过程中是否存在需要解决的异常情况。在更复杂的场景下，感知器应能够支持多模态（如文本、视觉和音频），以捕捉用户和环境反馈的多样性。

2.1 大模型工具学习流程

大模型工具学习的流程通常包括以下几个步骤。首先，用户提供指令，控制器需要根据指令制定一个可执行的计划。为了将用户指令与适当的工具连接起来，为了将用户指令与适当的工具连接起来，模型首先需要学习理解指令背后的用户意图，在此基础上，将复杂任务分解为若干子任务，并有效地用适当的工具完成每个子任务。在执行过程中，感知器会处理用户和环境反馈，并提供给控制器，以帮助调整和优化计划。

假设我们有一个工具集$\mathcal{T}$，控制器可以利用该工具集来完成某些任务。在执行的第t步，环境$\mathcal{E}$提供工具执行的反馈e_t。感知器$\mathcal{P}$接收用户反馈f_t和环境反馈e_t，生成总结反馈x_t。通常，感知器可以通过预定义规则（例如连接f_t和e_t）生成x_t，也可以使用复杂的神经模型进行建模。控制器$\mathcal{C}$生成当前步需要执行的计划a_t，选择并执行适当的工具进行执行。这个过程可以表示为以下概率分布：

$$p_{\mathcal{C}}(a_t) = p_{\theta_{\mathcal{C}}}(a_t \mid x_t, \mathcal{H}_t, q), \tag{1}$$

其中$\theta_{\mathcal{C}}$表示控制器的参数，q表示用户查询或指令，$\mathcal{H}t = (x_s, a_s)s = 0^{t-1}$表示历史反馈和计划。在其最简单的形式中，生成的计划a_t可以是工具执行的具体动作。控制器还可以将推理过程与动作预测相结合，a_t可能包含解释下一步应解决的子任务和选择解决该子任务的工具的推理轨迹。

在生成计划a_t之后，它将在环境$\mathcal{E}$中执行，环境反馈e_{t+1}将传递给感知器。上述过程重复进行，直到控制器完成任务。总体目标是找到一个动作序列a_t，最终实现用户指令q指定的任务。需要注意的是，在工具执行之后，控制器还可以将执行结果整合成一个合理的用户响应。

在理解和应用大模型工具学习的过程中，我们可以看到以下四个关键问题：

- **意图理解（Intent Understanding）**：理解用户任务意图是工具学习的首要步骤。控制器需要准确解析用户的输入，理解用户希望达成的目标和要求。这不仅仅是对用户文字表述的理解，还涉及到对隐含意图的推测和把握。例如，当用户询问“如何减肥”时，控制器需要明白用户不仅仅是在寻找一般性的建议，而可能需要具体的饮食计划、锻炼方案以及科学的减肥方法。

- **规划（Planning）**：在明确用户意图之后，控制器需要将用户的任务分解为一系列可执行的子任务。这一过程需要强大的规划和推理能力，以确保每个子任务都能有效地推动最终目标的实现。比如，对于“我想预订下周去北京的航班”这一任务，控制器需要分解出多个子任务，如查询航班信息、选择合适的航班、填写个人信息以及完成支付等步骤。

- **工具使用（Tool Use）**：一旦任务被分解为子任务，控制器需要选择和使用适当的工具来完成每个子任务。这要求控制器对可用工具的功能有深刻的理解，并能够根据任务需求灵活应用这些工具。例如，查询航班信息时可能需要调用航空公司API，而完成支付则需要调用支付网关的API。

- **记忆管理（Memory Management）**：在整个任务执行过程中，管理工作历史是确保任务顺利进行的重要因素。控制器需要跟踪每个子任务的执行状态，保存中间结果，并在需要时回顾和利用这些历史信息来调整和优化后续操作。有效的记忆管理能够帮助控制器在复杂和长时间的任务中保持高效和准确性。例如，在多轮对话中，控制器需要记住用户之前提供的信息和系统的响应，以便在后续对话中进行参考和调整。

在本文接下来的部分，我们将通过典型工作介绍在大模型工具学习中是如何处理意图理解、规划、工具使用、记忆管理问题的。具体来说，我们将探讨最新研究在这些领域所采取的方法和取得的进展，展示基础模型在实际应用中的表现和潜力。我们会深入分析每个问题的挑战，并探讨现有解决方案的有效性和局限性，最终为未来的研究和应用提供有价值的见解。

3 意图理解

理解用户意图一直是自然语言处理领域的一个长期研究课题 (Jansen et al., 2007; Sukthankar et al., 2014)。通过准确识别用户意图，控制器可以提供更加个性化的响应，从而提升

用户体验。近年来，指令调优（instruction tuning）方面的探索表明，基础模型在理解用户指令方面展现出了非凡的能力 (Wei et al., 2022a)。已有研究表明，将大语言模型在包含人类指令的多个数据集上进行微调，可以使模型甚至对未见过的任务指令进行泛化 (Wei et al., 2022a; Mishra et al., 2022; Sanh et al., 2022; Ouyang et al., 2022)。令人鼓舞的是，通过扩大模型规模以及增加训练指令的数量和多样性，可以进一步增强这种泛化能力 (Iyer et al., 2022)。

尽管意图理解能力已经取得了显著进展，但在实际的工具学习场景中仍存在一些挑战：

- **理解模糊指令（Understanding Vague Instructions）**：许多用户查询本质上是不精确的，甚至可能具有多义性，这要求控制器依赖上下文线索和背景知识来推断用户的真实意图。为了更好地理解用户意图，模型需要具备动态互动的能力。当遇到模糊或多义的指令时，模型应能够主动向用户询问澄清问题，以获取更多信息。这种互动不仅有助于提高指令理解的准确性，还能增强用户体验，使用户感受到被重视和理解。

  针对这个问题，钱等人的工作(Qian et al., 2024)介绍了一种创新的方法来解决大模型工具学习系统在处理用户指令时遇到的用户意图模糊问题。研究者们首先创建了一个名为Intention-in-Interaction (IN3)的新基准测试，它包含了一系列设计用来评估代理理解用户隐含意图能力的任务。这些任务被标注了不同程度的模糊性和缺失的细节，从而为评估提供了量化基础。接着，研究者们提出了将模型专家集成到大模型工具学习系统设计中的上游，以增强用户与代理之间的交互。特别是，他们开发了一个名为MistralInteract的模型，该模型通过模拟与用户的对话来明确任务的模糊性，主动询问缺失的细节，并在执行具体任务之前，将用户的意图明确化和细化为可操作的目标。为了训练MistralInteract，研究者们利用IN3的标注数据构建了模拟的对话记录，这些对话记录指导模型如何进行有效的交互。在实验中，MistralInteract被集成到了XAgent框架中，并通过一系列全面的评估来证明其在理解用户指令和执行任务方面的有效性。结果表明，MistralInteract在识别模糊任务、恢复关键信息、设置精确的执行目标以及减少不必要的工具使用方面表现出色，从而提升了整体的执行效率。这项工作不仅展示了如何通过用户参与来提高智能代理的性能，而且还通过开源数据和代码，为未来在这一领域的研究提供了基础和启发。通过这种方法，研究者们为构建更加用户友好的大模型工具学习系统迈出了重要的一步。

- **泛化到多样化指令（Generalization to Diverse Instructions）**：由于意图空间在理论上是无限的，因此基础模型在训练期间几乎不可能接触到所有现实世界中的意图表达。此外，个性化的挑战在于每个人都有自己独特的表达意图的方式，这要求模型适应不同个体的多样化意图表达。一种解决方案是利用用户反馈，主动适应个体用户，即个性化工具学习。

  由于大模型通常在通用领域进行训练，并根据广泛定义的人类偏好进行校准，这些偏好优先考虑有用性和无害性 (Ouyang et al., 2022; Nakano et al., 2021)。因此，它们在处理个人信息和提供个性化辅助方面存在困难。近年来，用户中心和个性化的自然语言生成受到了越来越多的关注 (Yang and Flek, 2021; Kirk et al., 2023)。现有工作涵盖了广泛的任务，如对话生成 (Madotto et al., 2019; Mazaré et al., 2018; Song et al., 2021; Zhong et al., 2022)、机器翻译 (Mirkin and Meunier, 2015; Michel and Neubig, 2018; Wuebker et al., 2018)和摘要生成 (Yan et al., 2011)。这些方法利用外部用户特定模块，如用户嵌入和用户记忆模块 (Zhang et al., 2018; Wu et al., 2021)，将不同用户的偏好、写作风格和个人信息注入生成的内容中。然而，这些工作通常针对特定任务设计，并在有限的用户信息下进行实验。如何将用户信息整合到大模型工具学习系统中仍然是一个未充分探索的领域。

  具体地，个性化工具学习强调在工具操作中考虑用户特定信息的重要性，主要有几个挑战：（1）异构用户信息建模：在现实世界中，个人信息可以来自多种异构来源。例如，使用电子邮件工具时，模型需要考虑用户历史对话记录中的语言风格，并从用户的社交网络中收集相关信息。这要求将具有多样结构的用户信息建模为统一的语义空间，允许模型联合利用这些信息。（2）个性化工具规划：不同用户在工具规划和选择上有不同的偏好。例如，在完成购买任务时，不同用户倾向于使用不同的在线购物平台。因此，模型需要根据用户偏好制定个性化的工具执行计划。（3）个性化工具调用：根据用户偏好自适应调用工具也是个性化工具学习中的重要方向。大多数工具在设计时未考虑个性化信息，这要求模

型根据用户偏好生成不同的工具输入。通过解决这些挑战，我们可以提高工具学习系统在理解用户意图和提供个性化支持方面的能力，从而提升整体用户体验。

4 规划

在大模型工具学习中，用户查询q通常涉及复杂任务，需要将其拆分为多个子任务并正确排序，这就需要强大的规划能力。最新研究表明，当预训练模型参数量达到一定规模时，其推理规划能力会显著增强 (Wei et al., 2022b)。拥有数百亿参数的预训练模型在解决复杂问题时，能够生成中间推理步骤，从而显著提升零样本和小样本任务的性能 (Nakano et al., 2021; Nye et al., 2021; Wei et al., 2022b)。然而，传统的少样本提示学习在处理需要复杂推理的问题时表现有限 (Creswell et al., 2022)。为此，(Wei et al., 2022c)提出了思维链推理（Chain-of-Thought，CoT）方法，通过在提示中插入推理步骤，引导模型生成解决问题的中间步骤，从而提升任务性能 (Wei et al., 2022c)。

基于大模型的强大推理规划能力，研究人员成功地将其应用于大模型工具学习的控制器中。其推理能力使控制器能够将复杂问题有效地分解为多个子问题，并确定每个子问题所需的工具。在这一方面，典型的研究工作主要分为两类：无反馈推理（*Planing without Feedback*）和带反馈推理（*Planning with Feedback*）。前者不与环境$\mathcal{E}$交互，生成静态任务拆解和工具调用计划；后者通过迭代地与环境$\mathcal{E}$交互，利用反馈逐步生成任务拆解和工具调用计划。

4.1 无反馈推理

无反馈推理是指在不依赖中间执行结果的情况下，直接生成多步计划。这种方法的一个典型例子是程序辅助语言模型（Program-Aided Language Models, PAL）(Gao et al., 2022)，该模型通过生成Python代码作为中间推理步骤，显著提高了大模型在算术、符号和算法方面的推理能力。PAL利用Python程序解释器作为工具，使模型能够像程序员一样编写详细注释，并在解决复杂问题时展示出强大的推理和规划能力。这一思路在具身自主智能体（embodied agents）中也得到了验证，如ProgPrompt(Singh et al., 2022)和Code-as-Policies (Liang et al., 2022)，这些方法通过生成可执行程序来指导具身自主智能体的实际行动，展示了模型在不直接与环境互动的情况下生成有效计划的能力。

另一无反馈推理的例子是Visual ChatGPT (Wu et al., 2023)，该系统将各种视觉预训练模型与ChatGPT结合，使其能够理解和生成图像。在Visual ChatGPT系统中，ChatGPT作为核心控制器，进行顺序决策。在每一步，ChatGPT可能调用一个视觉模型来修改现有图像或用纯文本回应用户。尽管这些模型没有直接与环境交互，但它们能够生成合理的中间步骤，并有效地处理复杂任务。

然而，无反馈推理方法的一个主要缺点是可能生成不切实际的计划。由于缺乏环境反馈，模型可能会在执行过程中遇到无法预见的异常情况。为了解决这一问题，SayCan (Ahn et al., 2022)提出了一种方法，通过使用价值函数估计每个动作成功执行的概率，使代理的行动更符合实际环境的约束。通过这种方式，工具学习系统能够在规划过程中考虑环境的实际情况，从而生成更为现实和可行的计划。

尽管存在这些挑战，无反馈推理方法在许多任务中仍然展示了其强大的规划能力。通过在生成计划时预见可能的异常情况并进行相应调整，无反馈推理方法能够在没有直接环境交互的情况下有效地解决复杂问题。

4.2 带反馈推理

相比之下，带反馈推理方法通过将环境$\mathcal{E}$纳入规划过程，逐步生成计划，使模型能够根据中间执行结果进行调整。这种方法更为灵活和动态，适合处理复杂任务，如多步问答和具体化学习，每一步的决策都依赖于之前的上下文。

带反馈推理的一个典型例子是Self-Ask (Press et al., 2022)、ReAct (Yao et al., 2022b)和ToolFormer (Schick et al., 2023)。这些研究表明，通过提供搜索引擎API访问权限，模型能够在多步问答中的准确性得到显著提高。通过思维链推理（CoT）提示或微调，这些模型能够将复杂问题分解为多个子问题，并利用搜索API找到每个子问题的答案。在获得每个子问题的答案后，模型能够迭代地确定下一个要问的问题或给出最终答案。

在具身学习中，带反馈推理方法通过直接与环境交互，进一步增强了模型的规划能力。例如，Inner Monologue (Huang et al., 2022)通过利用来自环境的多种反馈，如任务是否成功完成

和当前场景信息，生成更可行的计划，并提高高层次指令的完成能力。LLM-Planner (Song et al., 2022)明确考虑计划执行过程中可能出现的异常，并利用环境反馈在执行失败时重新生成计划，使模型能够适当处理例外情况。

此外，ReAct (Yao et al., 2022b)赋予模型自主权，使其在规划过程中能够根据当前情况决定何时停止生成动作令牌，从而制定更精细的后续计划。这种方法不仅提高了模型的灵活性，还增强了其在复杂任务中的表现能力。

尽管无反馈推理和带反馈推理各有优势，但将两者结合起来可以实现更强大的规划能力。在实际应用中，可以首先通过无反馈推理生成初步计划，然后利用带反馈推理在执行过程中不断调整和优化计划。例如，在复杂的任务如机器人控制和智能助手中，初步计划可以由内省推理生成，而具体执行时则通过外部推理实时调整，以应对环境中的变化和不确定性。这一思想被现在效果最好的大模型工具学习系统之一的XAgent所采用，使其能够真正处理需要几十步甚至上百步的复杂问题。

5 工具使用

在大模型工具学习中，工具使用指的是模型根据任务需要，选择并正确操作各种工具，以完成用户指令。工具使用的有效性直接影响任务的完成质量，因此，如何训练模型有效地使用工具至关重要。在这部分，我们将探讨如何通过不同的训练策略提升模型的工具使用能力。人类使用新工具主要有三种方式：通过学习其他人的示范、通过阅读工具手册或依靠自身探索尝试。同样地，我们将工具学习的训练策略分为三类：（1）示范学习：从具体的工具使用示范中学习，通常需要人工标注；（2）教程学习：从工具手册中学习；（3）探索学习：从工具探索尝试得到反馈中学习，通常涉及强化学习。

5.1 示范学习

从示范中学习是让模型通过观察和模仿人类专家的操作来掌握工具使用的方法。通过模仿学习，模型可以学到在特定情境下应如何使用工具。具体来说，行为克隆（Behavior Cloning）是一种常见的模仿学习形式，它假设人类专家的行为是最优或接近最优的，并通过监督学习来训练模型模仿这些行为。

假设我们有一个数据集$\mathcal{D}$，包含用户查询q和人类示范标注a^*的对，每对数据(q_i, a^i)表示一个具体的任务及其解决方案。我们的目标是优化控制器参数$\theta\mathcal{C}$，使其能够模仿人类专家的操作。学习目标可以表示为：

$$\theta_{\mathcal{C}}^* = \arg\max_{\theta_{\mathcal{C}}} \mathbb{E}_{(q_i,a_i^*)\in\mathcal{D}} \prod_{t=0}^{T_i} p_{\theta_{\mathcal{C}}}(a_{i,t}^* \mid x_{i,t}, \mathcal{H}_{i,t}, q_i), \tag{2}$$

其中$a_{i,t}^*$表示第i个任务执行到第t步时的人类标注。

具体来说，示范学习可分为三种主要方式：

- **监督学习（Supervised Learning）**：监督学习是最常见的示范学习形式，一般通过大量人类标注的数据来训练模型让其学会使用特定工具。例如，WebGPT (Nakano et al., 2021)通过与搜索引擎的交互，记录并模仿人类的搜索行为，来提升其信息检索能力。在这个过程中，研究人员首先建立了一个由Bing支持的搜索接口进行数据标注，然后采用标注出来的人类使用搜索引擎回答问题的行为序列来微调GPT-3，使其能够模仿人类专家的搜索行为。通过这种方式，WebGPT不仅能够生成有效的搜索查询，还能记录并总结重要的信息，从而提供更高质量的答案。另一个例子是WebShop (Yao et al., 2022a)，该模型在一个虚拟购物环境中学习如何根据人类指令进行商品购买。研究人员首先创建了一个互动环境，让模型能够浏览网页并选择商品，然后通过行为克隆训练模型模仿人类的购物行为，最终使模型能够在给定指令的情况下正确选择商品并完成购买。

- **半监督学习（Semi-supervised Learning）**：在很多情况下，获取大量高质量的人类标注数据是困难的，因此半监督学习提供了一种解决方案，即利用未标注数据生成伪标签，然后用这些伪标签来训练模型。例如，Baker等人的VPT工作(Baker et al., 2022)使用少量标注数据训练模型预测Minecraft视频游戏中每个时间步的动作伪标签，从而在没有大规模

人类行为标注数据的情况下，训练出更强大的模型。这种方法的核心在于利用少量的种子数据，训练一个初步模型来生成伪标签，再用这些伪标签进行更大规模的模型训练，从而在减少标注成本的情况下，提高模型的性能。

- **自监督学习（Self-supervised Learning）**：自监督学习进一步减少了对人工标注的依赖，模型通过自身的反馈迭代提升。例如，Toolformer工作 (Schick et al., 2023)利用少量人类示范，自动化生成工具使用示例，并通过过滤减少噪音，显著提升了工具使用性能。在这种方法中，模型首先使用一些基础示范进行初步学习，然后通过生成和筛选新的示范数据，不断改进自身的工具使用能力。这种方法的优势在于它能够利用现有的少量数据，通过自我学习和改进，逐步提升模型的性能。

5.2 教材学习

教程学习通过提示阅读工具手册，帮助模型理解工具的功能和使用方法。人类在学习使用新工具时，通常会通过阅读手册或观察他人演示来获取相关知识和技能。同样，模型也可以通过提示学习工具的使用方法。

在实际场景中，工具通常附带有使用手册（或教程），提供了关于其功能和用法的详细信息。大模型具备强大的零次学习（zero-shot learning）和少次学习（few-shot learning）能力，可以通过提示理解工具的功能和使用方法。具体来说，可以通过手动设计或检索构建合适的任务特定提示，这些提示描述了API的功能或通过示例展示其用法。

我们将提示方法分为两类：

- **零次提示（Zero-shot Prompting）**：描述API功能、输入输出格式、可能的参数等。这种方法使模型能够理解每个API可以处理的任务。

- **少次提示（Few-shot Prompting）**：为模型提供具体的工具使用示例。通过模仿这些示例中的人类行为，模型可以学习如何使用这些工具。

虽然提示方法具有显著优势，但也面临一些挑战。首先，提示受到输入上下文长度的限制。尽管大模型已经显示出通过提示学习使用简单工具的能力，但在面对多个复杂工具及其长描述时情况可能会更具挑战性。特别是当工具集大幅扩展时，在提示中提供所有可能的工具变得不可行，给定的上下文长度也有限。其次，也是最重要的一点，提示的效果很大程度上依赖于模型本身，较小或能力较弱的模型可能无法很好地理解提示。特别对于工具手册类的提示，几乎只有OpenAI系列大模型如ChatGPT、GPT-4拥有较强的零次提示、少次提示能力。

为了弥补这一差距，ToolLLM(Qin et al., 2023b)首先创建了一个名为ToolBench的指令调整数据集，它通过以下三个阶段自动构建：

1. API收集：从RapidAPI Hub收集了16,464个真实世界的RESTful APIs，覆盖了49个不同的类别。

2. 指令生成：使用ChatGPT生成涉及这些APIs的多样化指令，包括单工具和多工具场景。

3. 解决方案路径标注：利用ChatGPT为每个指令搜索有效的解决方案路径，即API调用链。

为了解决标注过程中大模型规划能力不足的问题，研究者们开发了一种新颖的基于深度优先搜索的决策树算法（DFSDT），它允许模型评估多种推理路径并扩展搜索空间。

基于ToolBench，研究者们微调了LLaMA模型，构建了ToolLLaMA模型，并为其配备了一个基于深度神经网络的API检索器，以推荐适合每个用户指令的APIs。实验结果表明，ToolLLaMA不仅能够执行复杂的指令，还能泛化到未见过的APIs，展现出与ChatGPT相当的性能。ToolLLaMA还在APIBench数据集上表现出强大的零样本泛化能力，证明了其在未知APIs上的适应性和灵活性。

5.3 探索学习

除了直接从人类示范和阅读工具手册学习使用工具外，模型还可以通过直接探索工具使用进行学习。这种方法在实际环境中探索使用工具，利用环境或人类的反馈来优化模型的工具使用策略。探索学习可以描述为通过开放探索优化控制器参数$\theta_{\mathcal{C}}$：

$$\theta_{\mathcal{C}}^{*}=\arg\max_{\theta_{\mathcal{C}}}\mathbb{E}_{q_i\in Q}\mathbb{E}_{\{a_{i,t}\}_{t=0}^{T_i}\in p_{\theta_{\mathcal{C}}}}\left[R(\{a_{i,t}\}_{t=0}^{T_i})\right], \tag{3}$$

其中R是从反馈序列中估计的奖励，T_i表示处理q_i所需的迭代次数。

强化学习（Reinforcement Learning, RL）通过与环境的互动，基于反馈信号（如奖励）来优化模型的决策过程。在大模型工具学习中，强化学习将动作空间定义为工具集中的所有工具，模型学习选择适当的工具并执行正确的动作以最大化奖励信号。例如，在机器人抓取任务中，模型通过反复尝试和调整抓取策略来学习最佳的工具使用方法 (Levine et al., 2018)。

在大模型工具学习中，我们主要从两个方面获取反馈：环境反馈和人类反馈。

- **环境反馈**：环境反馈包括模型与环境互动后得到的结果。根据反馈的不同，可以分为结果反馈和中间反馈：（1）结果反馈是任务完成与否的最终反馈，评估模型的整体表现。例如，WebShop (Yao et al., 2022a)通过评估模型购买的产品与人类购买的产品的相似性来提供反馈；（2）中间反馈是动作触发的环境状态变化，通过观察这些变化，模型可以学习每个动作的有效性，从而更好地调整其行为。例如，在信息检索任务中，模型可以通过观察搜索结果页面的内容来判断搜索查询的有效性，并根据这些信息调整后续的查询策略。这种反馈提供了关于每次工具执行效果的详细及时信息，使模型能够在任务执行过程中不断改进。

- **人类反馈**：人类反馈可以是显式的，如通过评分系统直接评价模型的行为；也可以是隐式的，通过用户行为和与模型的互动来推导用户的满意度。尽管人类反馈准确且稳定，但获取成本较高，因此人类反馈强化（RLHF）(Christiano et al., 2017)被提出，通过模仿人类给出奖励，然后使用强化学习算法来优化策略。例如，WebGPT(Nakano et al., 2021)利用人类反馈指导策略模型，使其在长篇问答中表现更好。

6 记忆管理

记忆管理问题仍是现有大模型工具学习系统中较少探索的领域。尽管在模拟型的大模型自主智能体中已经有初步的基于长短期记忆和外部检索模块的探索 (Park et al., 2023)，现有的大模型工具学习系统大多仍然依赖于通过提示语句的形式将信息拼接到大模型前面，利用大模型本身的长文本建模能力进行处理。然而，这种方法忽略了记忆管理中的一些关键问题。在复杂任务中，记忆管理不仅涉及到如何高效存储和检索信息，还包括如何在任务执行过程中动态更新和利用这些信息。具体来说，记忆管理面临以下几个挑战：

- **信息的持久性和可访问性**：大模型在处理长文本时可能会遇到上下文窗口的限制，导致无法有效利用所有相关信息。虽然通过提示语句拼接可以在一定程度上缓解这一问题，但这种方法无法保证信息的持久性和随时可访问性。对于需要长期记忆的任务，如持续对话或跨会话任务，现有方法显得力不从心。

- **信息的组织和优先级管理**：在任务执行过程中，不同信息的重要性和优先级可能不同。现有的方法往往无法区分和组织这些信息，导致重要信息可能被淹没在大量无关信息中。有效的记忆管理需要能够动态调整信息的组织方式，确保高优先级的信息能够被快速检索和利用。

- **动态更新和一致性维护**：随着任务的推进，新的信息不断涌入，旧的信息可能需要更新或淘汰。这就需要一个高效的机制来动态更新记忆内容，并保证信息的一致性。简单的拼接提示语句无法实现这一点，容易导致信息不一致或冗余。

7 总结

本文系统地探讨了大模型工具学习的最新进展及其核心问题。作为人类智慧的延伸，工具极大地提升了生产力和效率。当前，如何赋予大模型以工具学习能力，成为人工智能领域的前沿研究课题。

我们介绍了一个大模型工具学习的通用框架，包含控制器、工具集、感知器和环境四个核心组成部分。控制器负责理解用户意图并制定执行计划；工具集提供完成任务所需的各种工具；感知器处理用户和环境的反馈信息；环境为工具的执行提供平台和反馈。

基于这一框架，本文详细讨论了大模型工具学习中的几个关键问题：（1）意图理解：尽管大模型在指令理解方面已有显著进展，但理解模糊指令和泛化到多样化指令仍是主要挑战。通过主动互动和个性化学习，可以提升模型的理解能力。（2）规划与推理：在规划与推理方面，无反馈推理和带反馈推理是两种主要方法。结合这两种方法，可以更有效地应对复杂任务，确保任务拆解和工具调用的合理性和有效性。（3）工具使用：通过示范学习、教程学习和探索学习三种训练策略，可以显著提升模型的工具使用能力。示范学习通过模仿人类操作来训练模型；教程学习利用手册提示帮助模型理解工具功能；探索学习通过与环境互动，利用反馈优化工具使用策略。（4）记忆管理：记忆管理是大模型工具学习中的关键问题之一。现有方法主要依赖于提示语句拼接，但在处理复杂任务时存在局限。

综上所述，大模型工具学习在提升人工智能系统智能性和适应性方面展现出巨大潜力。通过持续的优化和创新，我们有望开发出更智能、更高效的人工智能系统，推动技术进步和社会发展。

参考文献

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, et al. 2022. Do as i can, not as i say: Grounding language in robotic affordances. *ArXiv preprint*, abs/2204.01691.

Ilge Akkaya, Marcin Andrychowicz, Maciek Chociej, Mateusz Litwin, Bob McGrew, Arthur Petron, Alex Paino, Matthias Plappert, Glenn Powell, Raphael Ribas, et al. 2019. Solving rubik's cube with a robot hand. *ArXiv preprint*, abs/1910.07113.

Bowen Baker, Ilge Akkaya, Peter Zhokhov, Joost Huizinga, Jie Tang, Adrien Ecoffet, Brandon Houghton, Raul Sampedro, and Jeff Clune. 2022. Video pretraining (vpt): Learning to act by watching unlabeled online videos. *ArXiv preprint*, abs/2206.11795.

Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Antonia Creswell, Murray Shanahan, and Irina Higgins. 2022. Selection-inference: Exploiting large language models for interpretable logical reasoning. *ArXiv preprint*, abs/2205.09712.

Wafaa S El-Kassas, Cherif R Salama, Ahmed A Rafea, and Hoda K Mohamed. 2021. Automatic text summarization: A comprehensive survey. *Expert systems with applications*, 165:113679.

Luyu Gao, Aman Madaan, Shuyan Zhou, Uri Alon, et al. 2022. Pal: Program-aided language models. *ArXiv preprint*, abs/2211.10435.

Shen Gao, Zhengliang Shi, Minghang Zhu, Bowen Fang, Xin Xin, Pengjie Ren, Zhumin Chen, Jun Ma, and Zhaochun Ren. 2024. Confucius: Iterative tool learning from introspection feedback by easy-to-difficult curriculum. In *In Proceedings of 38th Conference on Artificial Intelligence (AAAI)*.

Kathleen R Gibson, Kathleen Rita Gibson, and Tim Ingold. 1993. *Tools, language and cognition in human evolution*. Cambridge University Press.

Nicklas Hansen, Rishabh Jangir, Yu Sun, Guillem Alenyà, Pieter Abbeel, et al. 2021. Self-supervised policy adaptation during deployment. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.

Wenlong Huang, Fei Xia, Ted Xiao, Harris Chan, Jacky Liang, Pete Florence, Andy Zeng, Jonathan Tompson, Igor Mordatch, Yevgen Chebotar, et al. 2022. Inner monologue: Embodied reasoning through planning with language models. *ArXiv preprint*, abs/2207.05608.

Srinivasan Iyer, Xi Victoria Lin, Ramakanth Pasunuru, Todor Mihaylov, Dániel Simig, Ping Yu, et al. 2022. Opt-iml: Scaling language model instruction meta learning through the lens of generalization. *ArXiv preprint*, abs/2212.12017.

Bernard J. Jansen, Danielle L. Booth, and Amanda Spink. 2007. Determining the user intent of web search engine queries. In Carey L. Williamson, Mary Ellen Zurko, Peter F. Patel-Schneider, and Prashant J. Shenoy, editors, *Proceedings of the 16th International Conference on World Wide Web, WWW 2007, Banff, Alberta, Canada, May 8-12, 2007*, pages 1149–1150. ACM.

Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12):1–38.

Ehud Karpas, Omri Abend, Yonatan Belinkov, Barak Lenz, Opher Lieber, Nir Ratner, Yoav Shoham, Hofit Bata, Yoav Levine, Kevin Leyton-Brown, et al. 2022. Mrkl systems: A modular, neuro-symbolic architecture that combines large language models, external knowledge sources and discrete reasoning. *arXiv preprint arXiv:2205.00445*.

Hannah Rose Kirk, Bertie Vidgen, Paul Röttger, and Scott A Hale. 2023. Personalisation within bounds: A risk taxonomy and policy framework for the alignment of large language models with personalised feedback. *ArXiv preprint*, abs/2303.05453.

Tom Kwiatkowski, Jennimaria Palomaki, Olivia Redfield, Michael Collins, Ankur Parikh, Chris Alberti, Danielle Epstein, Illia Polosukhin, Jacob Devlin, Kenton Lee, et al. 2019. Natural questions: a benchmark for question answering research. *Transactions of the Association for Computational Linguistics*, 7:453–466.

Sergey Levine, Peter Pastor, Alex Krizhevsky, Julian Ibarz, and Deirdre Quillen. 2018. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International journal of robotics research*, 37(4-5):421–436.

Jacky Liang, Wenlong Huang, Fei Xia, Peng Xu, Karol Hausman, Brian Ichter, Pete Florence, and Andy Zeng. 2022. Code as policies: Language model programs for embodied control. *ArXiv preprint*, abs/2209.07753.

Andrea Madotto, Zhaojiang Lin, Chien-Sheng Wu, and Pascale Fung. 2019. Personalizing dialogue agents via meta-learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5459, Florence, Italy. Association for Computational Linguistics.

Alex Mallen, Akari Asai, Victor Zhong, Rajarshi Das, Daniel Khashabi, and Hannaneh Hajishirzi. 2022. When not to trust language models: Investigating effectiveness of parametric and non-parametric memories. *arXiv preprint arXiv:2212.10511*.

Pierre-Emmanuel Mazaré, Samuel Humeau, Martin Raison, and Antoine Bordes. 2018. Training millions of personalized dialogue agents. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2775–2779, Brussels, Belgium. Association for Computational Linguistics.

Paul Michel and Graham Neubig. 2018. Extreme adaptation for personalized neural machine translation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 312–318, Melbourne, Australia. Association for Computational Linguistics.

Shachar Mirkin and Jean-Luc Meunier. 2015. Personalized machine translation: Predicting translational preferences. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2019–2025, Lisbon, Portugal. Association for Computational Linguistics.

Swaroop Mishra, Daniel Khashabi, Chitta Baral, and Hannaneh Hajishirzi. 2022. Cross-task generalization via natural language crowdsourcing instructions. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3470–3487, Dublin, Ireland. Association for Computational Linguistics.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

Reiichiro Nakano, Jacob Hilton, Suchir Balaji, Jeff Wu, Long Ouyang, Christina Kim, Christopher Hesse, Shantanu Jain, Vineet Kosaraju, William Saunders, et al. 2021. Webgpt: Browser-assisted question-answering with human feedback. *ArXiv preprint*, abs/2112.09332.

Maxwell Nye, Anders Johan Andreassen, Guy Gur-Ari, Henryk Michalewski, Jacob Austin, David Bieber, et al. 2021. Show your work: Scratchpads for intermediate computation with language models. *ArXiv preprint*, abs/2112.00114.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *ArXiv preprint*, abs/2203.02155.

Liangming Pan, Xiaobao Wu, Xinyuan Lu, Anh Tuan Luu, William Yang Wang, Min-Yen Kan, and Preslav Nakov. 2023. Fact-checking complex claims with program-guided reasoning. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6981–7004, Toronto, Canada, July. Association for Computational Linguistics.

Aaron Parisi, Yao Zhao, and Noah Fiedel. 2022. Talm: Tool augmented language models. *ArXiv preprint*, abs/2205.12255.

Joon Sung Park, Joseph C O'Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. *arXiv preprint arXiv:2304.03442*.

Dean A Pomerleau. 1988. Alvinn: An autonomous land vehicle in a neural network. *Advances in neural information processing systems*, 1.

Ofir Press, Muru Zhang, Sewon Min, Ludwig Schmidt, Noah A Smith, and Mike Lewis. 2022. Measuring and narrowing the compositionality gap in language models. *ArXiv preprint*, abs/2210.03350.

Cheng Qian, Bingxiang He, Zhong Zhuang, Jia Deng, Yujia Qin, Xin Cong, Yankai Lin, Zhong Zhang, Zhiyuan Liu, and Maosong Sun. 2024. Tell me more! towards implicit user intention understanding of language model driven agents. *arXiv preprint arXiv:2402.09205*.

Yujia Qin, Shengding Hu, Yankai Lin, Weize Chen, Ning Ding, Ganqu Cui, Zheni Zeng, Yufei Huang, Chaojun Xiao, Chi Han, et al. 2023a. Tool learning with foundation models. *arXiv preprint arXiv:2304.08354*.

Yujia Qin, Shihao Liang, Yining Ye, Kunlun Zhu, Lan Yan, Yaxi Lu, Yankai Lin, Xin Cong, Xiangru Tang, Bill Qian, et al. 2023b. Toolllm: Facilitating large language models to master 16000+ real-world apis. *arXiv preprint arXiv:2307.16789*.

Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Arun Raja, Manan Dey, M Saiful Bari, et al. 2022. Multitask prompted training enables zero-shot task generalization. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.

Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2023. Toolformer: Language models can teach themselves to use tools. *ArXiv preprint*, abs/2302.04761.

Robert W Shumaker, Kristina R Walkup, and Benjamin B Beck. 2011. *Animal tool behavior: the use and manufacture of tools by animals*. JHU Press.

Ishika Singh, Valts Blukis, Arsalan Mousavian, Ankit Goyal, Danfei Xu, Jonathan Tremblay, Dieter Fox, Jesse Thomason, and Animesh Garg. 2022. Progprompt: Generating situated robot task plans using large language models. *ArXiv preprint*, abs/2209.11302.

Haoyu Song, Yan Wang, Kaiyan Zhang, Wei-Nan Zhang, and Ting Liu. 2021. BoB: BERT over BERT for training persona-based dialogue models from limited personalized data. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 167–177, Online. Association for Computational Linguistics.

Chan Hee Song, Jiaman Wu, Clayton Washington, Brian M Sadler, Wei-Lun Chao, and Yu Su. 2022. Llm-planner: Few-shot grounded planning for embodied agents with large language models. *ArXiv preprint*, abs/2212.04088.

Gita Sukthankar, Christopher Geib, Hung Bui, et al. 2014. *Plan, activity, and intent recognition: Theory and practice*. Newnes.

Dídac Surís, Sachit Menon, and Carl Vondrick. 2023. Vipergpt: Visual inference via python execution for reasoning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 11888–11898.

Tu Vu, Mohit Iyyer, Xuezhi Wang, Noah Constant, Jerry Wei, Jason Wei, Chris Tar, Yun-Hsuan Sung, Denny Zhou, Quoc Le, et al. 2023. Freshllms: Refreshing large language models with search engine augmentation. *arXiv preprint arXiv:2310.03214*.

Xingyao Wang, Zihan Wang, Jiateng Liu, Yangyi Chen, Lifan Yuan, Hao Peng, and Heng Ji. 2024. Mint: Evaluating llms in multi-turn interaction with tools and language feedback. *In Proceedings of 12th International Conference on Learning Representations (ICLR)*.

Sherwood L Washburn. 1960. Tools and human evolution. *Scientific American*, 203(3):62–75.

Jason Wei, Maarten Bosma, Vincent Y. Zhao, Kelvin Guu, Adams Wei Yu, et al. 2022a. Finetuned language models are zero-shot learners. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.

Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. 2022b. Emergent abilities of large language models. *ArXiv preprint*, abs/2206.07682.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed Chi, Quoc Le, and Denny Zhou. 2022c. Chain of thought prompting elicits reasoning in large language models. *ArXiv preprint*, abs/2201.11903.

Yuwei Wu, Xuezhe Ma, and Diyi Yang. 2021. Personalized response generation via generative split memory network. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1956–1970, Online. Association for Computational Linguistics.

Chenfei Wu, Shengming Yin, Weizhen Qi, Xiaodong Wang, Zecheng Tang, and Nan Duan. 2023. Visual chatgpt: Talking, drawing and editing with visual foundation models. *ArXiv preprint*, abs/2303.04671.

Joern Wuebker, Patrick Simianer, and John DeNero. 2018. Compact personalized models for neural machine translation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 881–886, Brussels, Belgium. Association for Computational Linguistics.

Qiantong Xu, Fenglu Hong, Bo Li, Changran Hu, Zhengyu Chen, and Jian Zhang. 2023. On the tool manipulation capability of open-source large language models. *arXiv preprint arXiv:2305.16504*.

Rui Yan, Jian-Yun Nie, and Xiaoming Li. 2011. Summarize what you are interested in: An optimization framework for interactive personalized summarization. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 1342–1351, Edinburgh, Scotland, UK. Association for Computational Linguistics.

Diyi Yang and Lucie Flek. 2021. Towards user-centric text-to-text generation: A survey. In *Text, Speech, and Dialogue: 24th International Conference, TSD 2021, Olomouc, Czech Republic, September 6–9, 2021, Proceedings 24*, pages 3–22. Springer.

Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W Cohen, Ruslan Salakhutdinov, and Christopher D Manning. 2018. Hotpotqa: A dataset for diverse, explainable multi-hop question answering. *arXiv preprint arXiv:1809.09600*.

Shunyu Yao, Howard Chen, John Yang, and Karthik Narasimhan. 2022a. Webshop: Towards scalable real-world web interaction with grounded language agents. *ArXiv preprint*, abs/2207.01206.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. 2022b. React: Synergizing reasoning and acting in language models. *ArXiv preprint*, abs/2210.03629.

Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. Personalizing dialogue agents: I have a dog, do you have pets too? In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213, Melbourne, Australia. Association for Computational Linguistics.

Yue Zhang, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, et al. 2023. Siren's song in the ai ocean: a survey on hallucination in large language models. *arXiv preprint arXiv:2309.01219*.

Tianyi Zhang, Faisal Ladhak, Esin Durmus, Percy Liang, Kathleen McKeown, and Tatsunori B Hashimoto. 2024. Benchmarking large language models for news summarization. *Transactions of the Association for Computational Linguistics*, 12:39–57.

Yuyue Zhao, Jiancan Wu, Xiang Wang, Wei Tang, Dingxian Wang, and Maarten De Rijke. 2024. Let me do it for you: Towards llm empowered recommendation via tool learning. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

Hanxun Zhong, Zhicheng Dou, Yutao Zhu, Hongjin Qian, and Ji-Rong Wen. 2022. Less is more: Learning to refine dialogue history for personalized dialogue generation. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5808–5820, Seattle, United States. Association for Computational Linguistics.