基于大语言模型的自主智能体概述

陈旭
高瓴人工智能学院
中国人民大学
xu.chen@ruc.edu.cn

摘要

近年来，基于大语言模型的自主智能体受到了学术界和工业界的广泛关注，其关键在于利用大语言模型作为核心控制器，并设计相应的辅助模块增强智能体在动态环境中的演化和适应能力，从而提升智能体自主解决任务的能力。本文通过总结过去工作，抽象出智能体设计的通用范式，并讨论了大模型时代自主智能体能力提升的途径。我们还从个体拓展到系统，深入探讨了多自主智能体系统常见的交互机制和面临的重要问题。

关键词： 大语言模型 ； 自主智能体

A Survey on Large Language Model based Autonomous Agents

Xu Chen
Gaoling School of Artificial Intelligence
Renmin University of China
xu.chen@ruc.edu.cn

Abstract

In recent years, autonomous agents based on large language models (LLMs) have garnered widespread attention from academia and industry. The central approach involves using LLMs as the primary controllers, complemented by auxiliary modules that enhance the agents' capabilities for evolution and adaptation in dynamic environments, thus boosting their proficiency in autonomously addressing tasks. This paper reviews previous efforts, distills a universal paradigm for agent design, and explores methods to advance the capabilities of autonomous agents in the era of LLMs. Furthermore, we extend our discussion from individual agents to systems, examining the prevalent interaction mechanisms and critical challenges faced by multi-agent systems.

Keywords: Large language model , Autonomous agent

1 引言

自主智能体作为实现通用人工智能的潜在途径之一，核心价值在于其能够独立规划和执行复杂场景下的任务。传统的自主智能体相关研究或者通过逻辑规则和符号表示来封装知识并进行推理 (Sacerdoti, 1975)，或者结合传统强化学习方法 (Watkins, 1989; Rummery and Niranjan, 1994)和深度神经网络 (Tesauro, 1995; Li, 2017)来从环境互动中不断学习并改进智能体自身的行为策略。基于以上传统算法的自主智能体已经在一些特定任务场景下取得了显著的成果，例如AlphaGo (Silver et al., 2016)和DQN (Mnih et al., 2013)。然而，这些早期探索仍然大多利用简单的启发式策略在封闭环境中进行学习，这一过程与人类复杂且灵活的学习方式存在着明显不同。具体来说，人类的思维过程极为复杂，涉及多种认知功能和学习机制，这使其能够在不

第二十三届中国计算语言学大会论文集，第141页-第150页，太原，中国，2024年7月25日至28日。
卷2：前沿综述

断变化的环境中快速学习和适应。由于这种差异的存在，传统研究中的自主智能体在面对开放和无限制的环境时，往往难以具备人类水平的决策能力。

近期，大语言模型（Large Language Models, LLMs）的相关研究取得了显著进展，展现出了强大的自然语言理解能力和实现人类水平智能的巨大潜力 (Wang et al., 2024)。这种令人瞩目的表现主要得益于对广泛的训练数据集和庞大的模型参数的深入利用 (Brown et al., 2020)。基于此，一个快速成长的新领域开始显现，即利用LLMs作为核心控制单元来构建高级的自主智能体，旨在实现与人类相似的决策水平。在这一研究领域，研究者们已经开发了多种有希望的智能模型，其核心理念是赋予LLMs关键的人类特性，例如记忆和规划能力，以期它们能够高效地完成各种任务。然而，值得注意的是，当前基于大语言模型的自主智能体研究大多是独立发展的，对于这个快速增长且潜力巨大的领域，目前还缺少一个系统的综述和比较分析。

本文广泛而深入地探讨了基于LLMs的自主智能体研究。具体而言，研究聚焦于在构建自主智能体过程中的两大挑战：一是设计一种能够最大化LLMs潜能的自主智能体架构；二是探索如何提升自主智能体执行多样化特定任务的能力。简单来说，第一个挑战相当于为自主智能体打造一个“硬件”平台，而第二个挑战则着眼于为自主智能体提供“软件”支持。针对第一个挑战，本文提出了一个全面的设计框架，整合了众多现有研究中的多种模块设计，以确保其在各种场景下的适应性和可扩展性；至于第二个挑战，本文归纳了一套常用的策略，旨在赋予智能体特定的能力。接下来，本文将对这两个关键议题进行详尽的阐述。

2 基于LLMs的自主智能体的架构设计

LLMs的迅猛发展在对话和问答领域彰显了其非凡的潜力，然而要创建一个具有自主性的智能体，仅依靠问答功能是远远不够的。自主智能体需要能够独立地感知周遭环境，并且能够根据当前情况采取相应的行为，从环境中学习，以实现类似人类的自我进化。为了缩小传统LLMs和自主智能体之间的差异，一个至关重要的步骤是构建一个合理的智能体架构，这将有助于LLMs充分发挥其潜力。在这一领域，已有研究提出了多种模块来提升LLMs的能力。本文对这些研究成果进行整合，从设定模块、记忆模块、规划模块和行动模块四个方面对已有工作进行总结 (Wang et al., 2024)。设定模块的核心功能是确定智能体所扮演的角色。记忆与规划模块则将智能体置于一个动态变化的环境中，使其能够回顾过往行为并预测未来行动。行动模块则负责将智能体的决策转化为实际的输出。在这些模块中，设定模块对记忆和规划模块有着直接的影响，而这三个模块又共同决定了行动模块的表现。本文将对这些模块进行详细介绍。

2.1 设定模块

自主智能体在执行任务时，经常扮演如程序员、教师或领域专家等特定角色。智能体的设定模块是为了指导智能体根据任务来设定其角色，这些设定信息通常通过提示词的形式嵌入，以影响LLMs的具体行为表现。智能体的设定通常包含基础信息，例如年龄、性别、职业，以及反映智能体性格特征的心理属性，还有描述智能体之间互动关系的社交属性。智能体设定的具体内容很大程度上取决于应用场景的需求。例如，如果应用的目的是探索人类的认知过程，那么心理属性就变得特别重要。在确定了设定文件所需包含的信息类型后，下一步是为智能体制定具体的设定文件，这通常涉及以下策略：例如，为了打造具有不同性格特征的智能体，人们可能会用“你是一个开朗的人”或者“你是一个害羞的人”来定义智能体的性格。这种手工定制方法已被许多研究工作采用，用以明确智能体的设定资料。例如，Expertprompting(Xu et al., 2023)通过人工精心设计不同领域的专家设定信息来增强智能体在不同领域的能力。Character-LLM (Shao et al., 2023)人工选择了七位名人作为智能体扮演的对象，结合各自维基百科的资料设计了每个智能体的设定。总的来说，这种人工设定方法具有很高的灵活性，因为可以为智能体指定任何所需的设定信息。但是，当涉及大量智能体时，这种方法可能会变得非常耗时。为缓解该问题，人们可以用大模型生成设定。RecAgent (Wang et al., 2023)首先由人工提供少量背景信息（如年龄、性别、个性特征和电影喜好）作为智能体的示例设定，然后利用ChatGPT根据这些种子信息生成更多的智能体设定。SOTOPIA(Zhou et al., 2023)通过GPT-4生成了多种具有不同行为类型的智能体及他们之间的社交关系，以研究智能体在动态环境中的社交能力。LLMs驱动策略在处理大量智能体时可以显著节省时间，但可能会在生成档案的精确度上有所欠缺。同时，人们也可以考虑使用真实数据集。在这种策略中，智能体的设定是基于真实世界数据集的信息。通常，首先将数据集中关于真实个体的信息整理成自然语言提示，然

后使用这些提示来设定智能体的档案。例如，EconAgent(Li et al., 2023)使用真实的年龄和收入信息作为智能体的设定信息，进而通过智能体来模拟宏观经济行为。数据集映射策略能够精确地捕捉到真实人群的属性，使得智能体的行为更加有意义，更能反映现实世界的情况。

2.2 记忆模块

在智能体的架构设计中，记忆组件扮演着至关重要的角色，它负责存储智能体从环境中获取的感知信息，并利用这些信息来指导未来的决策和行动。记忆组件不仅帮助智能体积累经验、自我进化，而且确保了其行为的一致性、合理性和有效性。接下来，本文将详细探讨记忆模块的结构、格式和操作。从结构上讲，基于LLMs的自主智能体通常借鉴人类认知科学的理论和方法。人类的记忆过程大致可以分为三个阶段：首先是感觉记忆，它负责存储从感官接收到的原始信息；其次是短期记忆，它用于暂时保留信息；最后是长期记忆，它负责将信息固定下来，以便长期保存。在设计这些智能体的记忆系统时，研究者们从人类记忆的这些阶段中汲取灵感。具体来说，短期记忆可以类比为受限于Transformer架构的上下文窗口，它能够处理并存储输入的信息；而长期记忆则类似于一个外部的存储系统，智能体可以迅速地访问和检索所需的信息。除了记忆结构之外，分析记忆模块的另一个角度是基于记忆存储介质的格式，例如自然语言记忆或编码向量记忆。不同的存储器格式具有不同的优势并且适合不同的应用。下面本文介绍几种有代表性的记忆格式。在现有研究中，Memochat(Lu et al., 2023)通过自我编写自然语言形式的备忘录来实现长期开放领域对话；Memory Sandbox(Huang et al., 2023)通过自然语言的形式存储记忆，便于用于对智能体的记忆进行管理，减少冗余记忆的干扰。MemoryBank (Zhong et al., 2023)通过将记忆片段转换为嵌入向量，构建了一个索引化的语料库，从而优化了记忆的检索过程。TiM(Liu et al., 2023)从原始信息中抽取出实体之间的关系作为记忆存储至数据库中来增强推理能力；在DB-GPT (Zhou et al., 2023)中，记忆模块同样也是基于数据库构建的，为了更直观地操作记忆信息，智能体经过微调以理解和执行SQL查询，使其能够直接使用自然语言与数据库进行交互。TradingGPT(Li et al., 2023)设计了分层记忆架构来存储不同的市场信息。从操作上讲，记忆模块在允许智能体通过与环境交互来获取、积累和利用重要知识方面发挥着关键作用。智能体与环境之间的交互是通过三个关键的记忆操作来完成的：记忆读取、记忆写入和记忆反思。记忆读取的目的是从记忆中提取有意义的信息以增强智能体的行动。例如，利用以前成功的行动来实现类似的目标。记忆读取的关键在于如何提取有价值的信息。通常，信息提取常用三个标准，即临近度、相关性和重要性 (Park et al., 2023)。记忆写入的目的是将从环境中感知到的有价值的信息存储在记忆中，这为将来检索信息中的丰富记忆奠定了基础，使智能体能够更高效、更合理地行动。在记忆写入的过程中，有两个潜在问题需要仔细解决。一方面，解决如何存储与现有记忆相似的信息（即记忆冗余）至关重要。另一方面，重要的是要考虑当记忆达到其存储限制（即记忆溢出）时如何删除信息。记忆反思旨在模仿人类见证和评估自己的认知、情感和行为过程的能力。当应用于智能体时，其目标是为智能体提供独立总结和推断更抽象、复杂和高级信息的能力。传统LLMs和智能体之间的一个显著区别是，后者必须具备在动态环境中学习和完成任务的能力。如果说记忆模块负责管理智能体过去的行为，那么就必须有另一个重要的规划模块来帮助智能体计划他们未来的行动。

2.3 规划模块

当面对复杂的任务时，人类会倾向于将其分解为更简单的子任务，并逐步解决；当人类在遭遇失败时，会对过去的错误原因进行反思，从失败中吸取教训。规划模块旨在赋予智能体类似的能力，从而使其可以解决复杂任务时更加有效与可靠。例如，在结构化思维链 (Li et al., 2023)中，作者用结构化的思维链约束大模型使用程序结构（例如：顺序、分支和循环结构）去组织思维过程，引导大模型从程序语言的角度去逐步思考如何解决需求。Q* (Wang et al., 2024)让大模型推理多条路径的同时，维护一个价值函数，综合考虑了当前状态的价值与未来期望价值，最后利用A*搜索算法对状态进行最佳优先搜索，实现了对复杂推理任务的全盘规划，在ToolLLM (Yao et al., 2023)中，智能体通过思考-调用工具-获取结果的循环来形成提示，其中思考环节用于促进深度的推理和策略规划，以指导智能体的行动；调用工具环节指的是智能体调用什么工具，怎么调用工具；而获取结果则是指动作产生的结果，这些结果通常是环境反馈的结果。此外，下一轮的思考过程会受到前一次观察的影响，从而使新的计划更加贴合环境条件。ReHAC(Feng et al., 2024)提出在agent多步规划的过程中，使用一个经过强化学习训练

后的策略模型，把子任务分配给人类，让经验丰富的人类来帮助基于LLMs的智能体更好地完成任务规划。在论文(Du et al., 2023)中，多个语言模型智能体在多轮中提出并辩论它们各自的回应和推理过程，以得出共同的最终答案。

2.4 行动模块

行动模块负责将智能体的决策转化为具体的结果。该模块位于最下游位置，直接与环境进行交互，受到设定、记忆和规划模块的影响。在目标方面，自主智能体的行动往往是为了完成各种目标，例如完成具有明确定义的任务（作为助手给用户推荐一家附近的餐厅 (Yao et al., 2023)）；与其他智能体或真实人类进行沟通，共享信息或合作 (Du et al., 2023)；探索陌生的环境以扩展感知,并在探索和利用之间取得平衡 (Wang et al., 2024)。在动作空间方面，动作空间指的是智能体可以执行的所有动作的集合。一般来说，本文可以将这些行动大致分为如下两类。外部工具：虽然LLMs已被证明在完成大量任务方面是有效的，但它们可能不适用于需要全面专业知识的领域。此外，LLMs也可能遇到难以自行解决的幻觉问题。为了缓解上述问题，智能体被赋予了调用外部工具（如API）来执行更多复杂操作的能力 (Yao et al., 2023)。内部知识：LLMs通常能够很好地理解人类常识并生成高质量的对话。基于这种能力，许多自主智能体可以直接利用其内在知识来模拟人类行为，从而做出类似人类的决策。

3 如何增强基于LLMs的自主智能体的特定能力

在上述内容中，本文详细探讨了基于LLMs的自主智能体的统一设计架构，旨在更好地激发LLMs的能力，使智能体能够以与人类相仿的智能水平完成复杂任务。自主智能体的架构相当于其“硬件”部分，但仅有硬件是不够的，因为智能体在面对特定任务时可能缺少必要的能力、技巧和知识，这些可以被看作是“软件”部分。为了弥补智能体在这些方面的不足，研究者们开发了多种策略。本文根据是否需要对LLMs进行微调，将这些策略划分为两大类，并将在下文中对每类策略进行详尽的阐述。

3.1 使用微调进行能力获取

提升自主智能体执行特定任务的效能，可以通过针对与任务相关的数据集进行微调来实现。这些数据集的来源可以多样化，包括但不限于人类专家的标注、由先进的语言模型生成的内容，或是从现实世界的应用场景中直接收集的数据。通过这样的微调，智能体能够更精准地理解和适应特定任务的需求，从而提高其性能。

利用人类注释数据集进行微调：微调LLMs以适应不同的应用场景，是一种通过人工标注数据集实现的高效策略。研究人员首先规划出需要的标注任务，随后招募人员来执行这些任务。例如，DPO (Liu et al., 2023)通过优化人类标注的偏好数据的对数似然，以实现LLMs与人类价值观的一致性，并直接利用这些自然语言数据集对模型进行微调;

利用LLMs生成的数据集进行微调：建立人类注释的数据集需要招募人员，这可能会带来高昂的成本，特别是当需要注释大量样本时。考虑到LLMs可以在广泛任务中实现类似人类的能力，一个自然的想法是使用LLMs来完成注释任务。虽然从这种方法产生的数据集可能不像人类注释的数据集那样完美，但它成本更低，并且可以用来生成更多的样本。例如，在FireAct (Chen et al., 2023)中，为了增强开源LLMs的推理能力，作者使用Chatgpt收集了一批专家数据，在复杂推理任务中获得了较好的性能，鲁棒性以及泛化性。

3.2 在没有微调的情况下获取能力

在基于LLMs的自主智能体时代，由于微调可能耗费大量资源，所以也可以采用非微调的方法提升模型能力。例如，在CoT (Wei et al., 2022)中，为了赋予智能体进行复杂任务推理的能力，作者将中间推理步骤作为少样本示例呈现在提示中。类似的技术也被用于CoT-SC (Wang et al., 2022)和ToT (Yao et al., 2024)中。EvoAgent(Yuan et al., 2024)通过进化算法，根据当前任务自动生成新的智能体从而自动化构建一个多智能体系统来解决当前任务。AgentHospital(Li et al., 2024)模拟了一家医院的运行流程，其中作为医生的智能体会根据历史信息积累经验，结合系统中已有的病历提高自己的诊断能力。ICE(Qian et al., 2024)将智能体的自我进化分解为探索，固化和利用三个阶段，将探索成功的路径保存为规划链，再下次执行任务时这些成功的规划链会作为参考的依据。

通过比较上述的自主智能体能力获取策略可以发现：微调策略通过优化模型参数，能够吸收特定任务的丰富知识，但这种方法主要适用于开源的LLMs；而无需微调的策略则依赖于精心设计的提示形式或机制工程来增强自主智能体的任务表现，这种策略同时适用于开源和闭源的LLMs，但由于LLMs对输入上下文的窗口大小有限制，它们往往无法处理过多的任务信息，同时提示和机制的设计空间非常大，这无疑增加了寻找最优解的难度。

4 由个体到系统：基于LLMs的多自主智能体系统

基于LLMs的单一智能体已经显示出了强大的认知能力，这些单一个体的开发主要集中在设计其内部工作方式及其对外部环境的响应。相比之下，由多个自主智能体组成的系统则强调每个智能体的角色属性、各个智能体间交互以及集体决策程序的多样化。多智能体系统旨在实现多个独立主体的协作，每个主体都被赋予了独特的策略和行为，促进彼此的沟通，进而有助于处理更动态、更复杂的任务。本节将深入探讨多智能体系统的内部交互形式和重要研究问题。

4.1 多智能体系统内部的交互方式

基于LLMs的多自主智能体系统旨在模拟人类群体动态，通过智能体之间的合作、竞争或层次化交互来解决复杂问题。智能体的交互方式对于整个系统的效能至关重要，它们可以是合作互补的，也可以是相互竞争的，甚至是动态变化的。

合作交互：合作型多智能体系统是实际应用中最广泛部署的模式。在这些系统中，每个智能体评估其对应智能体的需求和能力，积极追求协作行动并共享信息 (Li et al., 2023)。这种框架提供了多种潜在优势，例如提高任务效率、增强集体决策能力，以及解决单个智能体无法单独解决的复杂现实世界问题，最终实现整体系统性能的提升。SPP (Wang et al., 2023)通过利用多角色自我合作，实现了多轮对话，有效地将单一的大型语言模型转变为认知协同体；Generative Agents (Park et al., 2023)利用基于LLMs的多智能体系统模仿现实人类行为，促进智能体之间的合作；CAMEL (Li et al., 2023)通过任务导向的角色扮演实现AI助手与用户之间的合作多轮对话；MetaGPT (Hong et al., 2023)将高效的工作流程整合到LLMs驱动的多智能体协作编程方法中，促进了不同角色之间的协作；ChatDev (Qian et al., 2023)使用基于LLMs的多个智能体进行对话交流并解决任务，为加快软件应用的设计与开发提供了新思路；受Minsky心智社会概念 (Minsky , 1988)的启发，NLSOM (Zhuge et al., 2023)引入了基于自然语言的心智社会（NLSOMs）的概念，由多个LLMs和其他基于神经网络的专家通过自然语言界面进行通信。这种方法被应用于解决不同场景中的复杂任务；在具身智能领域，RoCo (Mandi et al., 2023)利用LLMs进行高层通信和低层路径规划，从而促进了多个机器人之间的协作；Interact (Chen and Chang , 2023)为各种角色（如检查员和分类器）的自主智能体分配任务，在AlfWorld (Shridhar et al., 2020)环境中取得了显著的成功率；AutoAgents (Chen et al., 2023)能够适应性地生成和协调多个专业智能体，从而形成一个强大的AI团队，在各种复杂任务场景中协作实现目标。

竞争交互：在竞争场景中，各个自主智能体往往发展出一系列策略和技能，例如制定稳健的行动计划和分析实时的竞争反馈，以确保自身能在激烈的彼此对抗中获得优势。在这种方式下，多智能体系统不仅能够通过内部竞争实现更加优越的整体表现，还能够在面对复杂挑战时展现出更高的适应性和创新能力。近期，Liang等人 (Liang et al., 2023)通过引入多智能体辩论框架增强了整体系统解决问题的能力；ChatEval (Chan et al., 2023)同样采用多智能体策略，使LLMs与多样化的智能对手进行互动，通过利用各个内部智能体的独特能力和专业知识，显著增强了系统解决复杂任务的效率和效果。

层次交互：层次型多智能体系统关注于在层次结构中开发高效控制结构、信息传输方法和任务分解技术。这些策略促进了不同层级智能体之间的有效协作，提高了整体系统性能。层次型多智能体系统通常以树状结构组织，父节点智能体负责任务分解并将任务分配给子节点智能体。后者遵循来自其父节点的指令并提供汇总反馈。例如，AutoGen (Wu et al., 2023)使用各种智能体执行代码生成和文本写作等任务，通过对话进行任务分解。关于层次型多智能体系统的研究仍然处于起步阶段，目前的探索仍局限于少数几个层级。

动态交互：动态互动指的是多智能体系统中结构的灵活性，其中智能体的角色、它们的关系以及智能体的总数可以随时间演变。动态交互型多智能体系统往往表现出强大的上下文适应性，即互动模式根据内部或外部影响进行调整，其中的智能体具有根据变化条件动态调整各自

角色和彼此之间关系的能力。例如，(Talebirad and Nadiri , 2023)展示了根据特定任务添加或删除智能体的可能性。

混合交互：多智能体系统中的混合交互需要在合作和竞争之间进行微妙的平衡以实现期望的结果。当前基于LLMs的多智能体系统研究聚焦于开发协作竞争算法，这是一个关键的研究领域。这种算法能够有效地帮助智能体系统在复杂的环境场景中做出更优的整体决策。近期，Xu等人 (Xu et al., 2023)使多个基于大语言模型的自主智能体参与狼人杀游戏，各个智能体根据不对称的信息来合作或背叛他人以实现各自的目标；Light (Light et al., 2023)等人则在阿瓦隆游戏场景下设计了一个多智能体系统，其中各个智能体导航动态发展的游戏阶段，并与其他智能体进行合作或欺骗以履行其指定角色；受人类行为启发，Corex (Sun et al., 2023)结合了辩论、审查和检索等多种合作范式，旨在增强推理过程的真实性、保真度和可靠性。

4.2 多智能体系统研究的开放问题

本小节对基于LLMs的多智能体系统中的开放问题进行探讨，旨在为这一迅速发展的研究领域指明进一步探索和创新的可行方向。

迈向多模态环境：近期关于基于LLMs的多智能体系统的研究主要集中在基于文本的环境上，展示了它们在文本处理和生成方面的专业能力。然而，在多模态设置的背景下仍然存在着一个显著的空白。在多模态环境中，多智能体系统往往需要处理来自各种感官输入的数据，并以多种模态产生输出，如图像、音频、视频和物理动作，其中的挑战主要在于处理和整合不同数据类型的复杂性，这要求智能体具有高级的感知和认知能力。此外，生成连贯且符合上下文的多模态输出要求各个智能体保持共享理解并有效协调它们的行动。解决这些挑战对于开发能够适应各种现实世界场景的多功能和适应性强的多智能体系统至关重要。

共同获取集体智能：传统的多智能体系统通常依赖于离线训练数据集的强化学习。相比之下，基于LLMs的多智能体系统主要利用通过与环境或人类交互获得的即时反馈。然而，这种学习范式需要一个可靠的交互环境，为跨各种任务的可扩展性带来了挑战。此外，当前研究中的主要方法强调使用记忆和自我演化技术，根据反馈适应个体智能体。尽管对于每个个体智能体有效，但这些方法未能充分利用多智能体系统的潜在集体智能。通过单独调整智能体，它们忽视了来自协调的多智能体交互的协同效应。因此，同时调整多个智能体以实现最佳集体智能仍然是基于LLMs的多智能体系统面临的关键挑战。

扩大智能体数量：增加智能体数量可以提高任务效率并增强社会仿真的真实性 (Park et al., 2023; Qian et al., 2023)。然而，现有研究仍然存在许多挑战。随着部署的人工智能智能体数量增加，计算负载也随之增加，这需要改进架构设计和计算优化以确保系统平稳运行。随着智能体数量的增长，通信和消息传播构成了重大障碍，导致高度复杂的通信网络。在多智能体系统中，由于幻觉和误解导致的信息传播偏见可能会扭曲信息传播，尤其是在智能体数量较多时，增加了风险并降低了通信的可靠性。此外，随着智能体数量的增加，协调智能体变得越来越困难，可能阻碍合作和效率，从而影响实现共享目标的进展。

5 结论

本文详细探讨了以LLMs为核心的自主智能体技术，主要聚焦于体系架构的设计,特定能力的获取和多自主智能体这三个关键议题。本文以一种统一的框架视角，系统地涵盖了当前主流的智能体构建技术，归纳总结了一系列赋予智能体特定能力的关键策略并分析了多自主智能体系统内部的交互机制和可能的开放问题。通过对现有相关工作的综合梳理，本文充分阐述了基于LLMs的自主智能体的研究现状和发展趋势。这些深入的讨论与思考不仅有助于读者更加全面地把握这一新兴技术领域，也为未来的探索和创新提供了宝贵的指导。随着技术的持续发展和理论的逐步深化，以LLMs为核心的智能体系统有望在各个领域发挥重要作用，为人类社会注入新的发展动力和创新潜力。

参考文献

J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pages 1–22.

S. Hong, X. Zheng, J. Chen, Y. Cheng, J. Wang, C. Zhang, Z. Wang, S. K. S. Yau, Z. Lin, L. Zhou, et al. 2023. Metagpt: Meta programming for multi-agent collaborative framework. *arXiv preprint arXiv:2308.00352*.

C. Qian, X. Cong, C. Yang, W. Chen, Y. Su, J. Xu, Z. Liu, and M. Sun. 2023. Communicative agents for software development. *arXiv preprint arXiv:2307.07924*.

Y. Dong, X. Jiang, Z. Jin, and G. Li. 2023. Self-collaboration code generation via chatgpt. *arXiv preprint arXiv:2304.07590*.

L. Wang, J. Zhang, X. Chen, Y. Lin, R. Song, W. X. Zhao, and J. R. Wen. 2023. Recagent: A novel simulation paradigm for recommender systems. *arXiv preprint arXiv:2306.02552*.

L. P. Argyle, E. C. Busby, N. Fulda, J. R. Gubler, C. Rytting, and D. Wingate. 2023. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351.

N. Shinn, F. Cassano, A. Gopinath, K. Narasimhan, and S. Yao. 2024. Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems*, 36.

G. Wang, Y. Xie, Y. Jiang, A. Mandlekar, C. Xiao, Y. Zhu, L. Fan, and A. Anandkumar. 2023. Voyager: An open-ended embodied agent with large language models. *arXiv preprint arXiv:2305.16291*.

W. Zhong, L. Guo, Q. Gao, and Y. Wang. 2023. Memorybank: Enhancing large language models with long-term memory. *arXiv preprint arXiv:2305.10250*.

X. Zhu, Y. Chen, H. Tian, C. Tao, W. Su, C. Yang, G. Huang, B. Li, L. Lu, X. Wang, et al. 2023. Ghost in the minecraft: Generally capable agents for open-world environments via large language models with text-based knowledge and memory. *arXiv preprint arXiv:2305.17144*.

C. Hu, J. Fu, C. Du, S. Luo, J. Zhao, and H. Zhao. 2023. Chatdb: Augmenting llms with databases as their symbolic memory. *arXiv preprint arXiv:2306.03901*.

X. Zhou, G. Li, and Z. Liu. 2023. Llm as dba. *arXiv preprint arXiv:2308.05481*.

A. Modarressi, A. Imani, M. Fayyaz, and H. Schütze. 2023. Retllm: Towards a general read-write memory for large language models. *arXiv preprint arXiv:2305.14322*.

J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, D. Zhou. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837.

X. Wang, J. Wei, D. Schuurmans, Q. Le, E. Chi, S. Narang, A. Chowdhery, and D. Zhou. 2022. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*.

S. Yao, D. Yu, J. Zhao, I. Shafran, T. Griffiths, Y. Cao, and K. Narasimhan. 2024. Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, 36.

B. Liu, Y. Jiang, X. Zhang, Q. Liu, S. Zhang, J. Biswas, and P. Stone. 2023. LLM+P: Empowering large language models with optimal planning proficiency. *arXiv preprint arXiv:2304.11477*.

S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan, and Y. Cao. 2023. React: Synergizing reasoning and acting in language models. In *The Twelfth International Conference on Learning Representations*.

W. Huang, F. Xia, T. Xiao, H. Chan, J. Liang, P. Florence, A. Zeng, J. Tompson, I. Mordatch, Y. Chebotar, et al. 2022. Inner monologue: Embodied reasoning through planning with language models. *arXiv preprint arXiv:2207.05608*.

A. Madaan, N. Tandon, P. Gupta, S. Hallinan, L. Gao, S. Wiegreffe, U. Alon, N. Dziri, S. Prabhumoye, Y. Yang, et al. 2024. Self-refine: Iterative refinement with selffeedback. *Advances in Neural Information Processing Systems*, 36.

Z. Wang, S. Cai, G. Chen, A. Liu, X. Ma, and Y. Liang. 2023. Describe, explain, plan and select: Interactive planning with large language models enables open-world multitask agents. *arXiv preprint arXiv:2302.01560*.

M. Ahn, A. Brohan, N. Brown, Y. Chebotar, O. Cortes, B. David, C. Finn, C. Fu, K. Gopalakrishnan, K. Hausman, et al. 2022. Do as i can, not as i say: Grounding language in robotic affordances. *arXiv preprint arXiv:2204.01691*.

H. Liu, C. Sferrazza, and P. Abbeel. 2023. Chain of hindsight aligns language models with feedback. In *The Twelfth International Conference on Learning Representations*.

B. Y. Lin, Y. Fu, K. Yang, F. Brahman, S. Huang, C. Bhagavatula, P. Ammanabrolu, Y. Choi, and X. Ren. 2024. Swiftsage: A generative agent with fast and slow thinking for complex interactive tasks. *Advances in Neural Information Processing Systems*, 36.

Y. Qin, S. Liang, Y. Ye, K. Zhu, L. Yan, Y. Lu, Y. Lin, X. Cong, X. Tang, B. Qian, et al. 2023. Toolllm: Facilitating large language models to master 16000+ real-world apis. *arXiv preprint arXiv:2307.16789*.

X. Deng, Y. Gu, B. Zheng, S. Chen, S. Stevens, B. Wang, H. Sun, and Y. Su. 2024. Mind2web: Towards a generalist agent for the web. *Advances in Neural Information Processing Systems*, 36.

K. A. Fischer. 2023. Reflective linguistic programming (rlp): A stepping stone in socially-aware agi (socialagi). *arXiv preprint arXiv:2305.12647*.

Y. Shu, H. Gu, P. Zhang, H. Zhang, T. Lu, D. Li, and N. Gu. 2023. Rah! recsys-assistant-human: A human-central recommendation framework with large language models. *arXiv preprint arXiv:2308.09904*.

Y. Du, S. Li, A. Torralba, J. B. Tenenbaum, and I. Mordatch. 2023. Improving factuality and reasoning in language models through multiagent debate. *arXiv preprint arXiv:2305.14325*.

C. Colas, L. Teodorescu, P. Y. Oudeyer, X. Yuan, and M. A. Côté. 2023. Augmenting autotelic agents with large language models. *arXiv preprint arXiv:2305.12487*.

N. Nascimento, P. Alencar, and D. Cowan. 2023. Self-adaptive large language model (llm)-based multiagent systems. In *2023 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C)*, pages 104–109.

S. Saha, P. Hase, and M. Bansal. 2023. Can language models teach weaker agents? teacher explanations improve students via theory of mind. *arXiv preprint arXiv:2306.09299*.

L. Wang, C. Ma, X. Feng, et al. 2024. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6): 1-26.

Chris Watkins. 1989. Learning from delayed rewards.

Gavin A. Rummery and Mahesan Niranjan. 1994. On-line Q-learning using connectionist systems. *University of Cambridge, Department of Engineering Cambridge, UK.*

Gerald Tesauro. 1995. Temporal difference learning and td-gammon. *Communications of the ACM*, 38(3): 58–68.

Yuxi Li. 2017. Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*.

D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. 2016. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587): 484–489.

V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, M. Riedmiller. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.

E. D. Sacerdoti. 1975. The nonlinear nature of plans. In *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence*, pages 206–214.

T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33: 1877–1901.

Kojima, Takeshi and Gu, Shixiang Shane and Reid, Machel and Matsuo, Yutaka and Iwasawa, Yusuke 2022. Language models are few-shot learners. *Advances in neural information processing systems*, 35: 22199–22213.

Chen, Baian and Shu, Chang and Shareghi, Ehsan and Collier, Nigel and Narasimhan, Karthik and Yao, Shunyu 2023. Fireact: Toward language agent fine-tuning. *arXiv preprint arXiv:2310.05915*.

G. Li, H. Hammoud, H. Itani, D. Khizbullin, and B. Ghanem. 2023. Camel: Communicative agents for "mind" exploration of large language model society. *Advances in Neural Information Processing Systems*, 36: 51991–52008.

Z. Wang, S. Mao, W. Wu, T. Ge, F. Wei, and H. Ji. 2023. Unleashing cognitive synergy in large language models: A task-solving agent through multi-persona self collaboration. *arXiv preprint arXiv:2307.05300*, 1(2): 3.

M. Minsky. 1988. Society of mind. *Simon and Schuster*.

M. Zhuge, H. Liu, F. Faccio, D. R. Ashley, R. Csordás, A. Gopalakrishnan, A. Hamdi, H. A. A. K. Hammoud, V. Herrmann, K. Irie, et al. 2023. Mindstorms in natural language-based societies of mind. *arXiv preprint arXiv:2305.17066*.

P.-L. Chen and C.-S. Chang. 2023. Interact: Exploring the potentials of chatgpt as a cooperative agent. *arXiv preprint arXiv:2308.01552*.

Z. Mandi, S. Jain, and S. Song. 2023. Roco: Dialectic multi-robot collaboration with large language models. *arXiv preprint arXiv:2307.04738*.

M. Shridhar, X. Yuan, M.-A. Côté, Y. Bisk, A. Trischler, and M. Hausknecht. 2020. Alfworld: Aligning text and embodied environments for interactive learning. *arXiv preprint arXiv:2010.03768*.

Q. Wu, G. Bansal, J. Zhang, Y. Wu, S. Zhang, E. Zhu, B. Li, L. Jiang, X. Zhang, and C. Wang. 2023. Autogen: Enabling next-gen llm applications via multi-agent conversation framework. *arXiv preprint arXiv:2308.08155*.

T. Liang, Z. He, W. Jiao, X. Wang, Y. Wang, R. Wang, Y. Yang, Z. Tu, and S. Shi. 2023. Encouraging divergent thinking in large language models through multi-agent debate. *arXiv preprint arXiv:2305.19118*.

J. Light, M. Cai, S. Shen, and Z. Hu. 2023. From text to tactic: Evaluating llms playing the game of avalon. *arXiv preprint arXiv:2310.05036*.

Q. Sun, Z. Yin, X. Li, Z. Wu, X. Qiu, and L. Kong. 2023. Corex: Pushing the boundaries of complex reasoning through multi-model collaboration. *arXiv preprint arXiv:2310.00280*.

Y. Talebirad and A. Nadiri. 2023. Multi-agent collaboration: Harnessing the power of intelligent llm agents. *arXiv preprint arXiv:2306.03314*.

Y. Xu, S. Wang, P. Li, F. Luo, X. Wang, W. Liu, and Y. Liu. 2023. Exploring large language models for communication games: An empirical study on werewolf. *arXiv preprint arXiv:2309.04658*.

G. Chen, S. Dong, Y. Shu, G. Zhang, S. Jaward, K. Börje, J. Fu, and Y. Shi. 2023. Autoagents: The automatic agents generation framework. *arXiv preprint*.

C.-M. Chan, W. Chen, Y. Su, J. Yu, W. Xue, S. Zhang, J. Fu, and Z. Liu. 2023. Chateval: Towards better llm-based evaluators through multi-agent debate. *arXiv preprint arXiv:2308.07201*.

Y. Shao, L. Li, J. Dai, and X. Qiu. 2023. Character-llm: A trainable agent for role-playing. *arXiv preprint arXiv:2310.10158*.

Chaojie Wang and Yanchen Deng and Zhiyi Lv and Zeng Liang and Jujie He and Shuicheng Yan and An Bo. 2024. Q*: Improving Multi-step Reasoning for LLMs with Deliberative Planning. *arXiv preprint arXiv:2406.14283*.

C. Shen, G. Xie, X. Zhang, and J. Xu. 2024. On the Decision-Making Abilities in Role-Playing using Large Language Models. *arXiv preprint arXiv:2402.18807*.

X. Zhou, H. Zhu, L. Mathur, R. Zhang, H. Yu, Z. Qi, L.-P. Morency, Y. Bisk, D. Fried, G. Neubig, *et al.* 2023. Sotopia: Interactive evaluation for social intelligence in language agents. *arXiv preprint arXiv:2310.11667*.

Xueyang Feng and Zhi-Yuan Chen and Yujia Qin and Yankai Lin and Xu Chen and Zhiyuan Liu and Ji-Rong Wen, *et al.* 2024. Large Language Model-based Human-Agent Collaboration for Complex Task Solving. *arXiv preprint arXiv:2402.12914*.

N. Li, C. Gao, Y. Li, and Q. Liao. 2023. Large language model-empowered agents for simulating macroeconomic activities. *arXiv preprint arXiv:2310.10436*.

Z. Huang, S. Gutierrez, H. Kamana, and S. MacNeil. 2023. Memory sandbox: Transparent and interactive memory management for conversational agents. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, pages 1–3.

L. Liu, X. Yang, Y. Shen, B. Hu, Z. Zhang, J. Gu, and G. Zhang. 2023. Think-in-memory: Recalling and post-thinking enable llms with long-term memory. *arXiv preprint arXiv:2311.08719*.

Jia Li and Ge Li and Yongmin Li and Zhi Jin 2023. Structured Chain-of-Thought Prompting for Code Generation. *arXiv preprint arXiv:2305.06599*.

Yilun Du and Shuang Li and Antonio Torralba and Joshua B. Tenenbaum and Igor Mordatch. 2023. Think-in-memory: Recalling and post-thinking enable llms with long-term memory. *arXiv preprint arXiv:2305.14325*.

Rafael Rafailov and Archit Sharma and Eric Mitchell and Stefano Ermon and Christopher D. Manning and Chelsea Finn. 2023. Direct Preference Optimization: Your Language Model is Secretly a Reward Model. *arXiv preprint arXiv:2305.18290*.

B. Xu, A. Yang, J. Lin, Q. Wang, C. Zhou, Y. Zhang, and Z. Mao. 2023. Expertprompting: Instructing large language models to be distinguished experts. *arXiv preprint arXiv:2305.14688*.

J. Lu, S. An, M. Lin, G. Pergola, Y. He, D. Yin, X. Sun, and Y. Wu. 2023. Memochat: Tuning LLMs to use memos for consistent long-range open-domain conversation. *arXiv preprint arXiv:2308.08239*.

Y. Li, Y. Yu, H. Li, Z. Chen, and K. Khashanah. 2023. TradingGPT: Multi-agent system with layered memory and distinct characters for enhanced financial trading performance. *arXiv preprint arXiv:2309.03736*.

J. Li, S. Wang, M. Zhang, W. Li, Y. Lai, X. Kang, M. Ma, and Y. Liu. 2024. Agent hospital: A simulacrum of hospital with evolvable medical agents. *arXiv preprint arXiv:2405.02957*.

C. Qian, S. Liang, Y. Qin, Y. Ye, X. Cong, Y. Lin, Y. Wu, Z. Liu, and M. Sun. 2024. Investigate-consolidate-exploit: A general strategy for inter-task agent self-evolution. *arXiv preprint arXiv:2401.13996*.

S. Yuan, K. Song, J. Chen, X. Tan, D. Li, and D. Yang. 2024. EvoAgent: Towards Automatic Multi-Agent Generation via Evolutionary Algorithms. *arXiv preprint arXiv:2406.14228*.