arXiv:2305.01253v1 [cs.CL] 2 May 2023

摘要在生成式智能体中的作用：初步观点

Xiachong Feng,Xiaocheng Feng,Bing Qin哈尔滨工业大学，中国 {xiachongfeng,xcfeng,bqin}@ir.hit.edu.cn

摘要

生成式智能体（Park 等人，2023年）能够模拟人类社会，在后续研究和实际应用中展现出巨大潜力。具体而言，由多个经过精心设计的模块组成的生成式智能体架构，是其中最关键的部分。为推动该领域的研究进展，本报告提出了我们对通过摘要理解生成式智能体的综合观点，因为我们认为，摘要是生成式智能体在各种场景中所体现出的最基本且不可或缺的能力。我们希望本报告能够帮助人们认识摘要能力对于生成式智能体的重要性，并推动未来研究。

1 引言

大语言模型（LLM），例如 ChatGPT 和 GPT-4（OpenAI，2023年），近期取得的进展重塑了多个领域，包括自然语言处理（Yang 等人，2023年）、计算机视觉（Wu 等人，2023年）和自主机器人学（Mai 等人，2023年）。这些前沿模型为实现通用人工智能（AGI）带来了新的机遇。随着大语言模型的快速发展，人们逐渐达成共识：大语言模型已初步展现出智能，并且如今在多种任务上的表现已可与人类媲美（Zhao 等人，2023年）。

在大语言模型时代，Park 等人（2023年）提出了生成式智能体：由基础语言模型驱动的复杂计算软件，能够在精心设计的环境和协议中模拟可信的人类行为。这一精心设计的框架为探索和理解人类社会动态提供了全面的机会，包括长期目标规划、信息转换、关系建立与协调。

在本报告中，我们从自动摘要的角度阐述对生成式智能体的看法，并展示如何将此类智能体的各种功能组件形式化为摘要任务。具体而言，我们归纳出实现生成式智能体不可或缺的若干关键摘要技术：（1）检索模块包含无监督摘要的思想（§3.1）；（2）反思模块由两个子模块构成：极端摘要（§3.2）和基于引文的摘要（§3.3）；（3）基于查询的摘要（§3.4）支持后续的规划模块和行动模块；（4）带表情符号的摘要（§3.5）提供直观的视觉界面；（5）智能体在环境中的移动可以抽象为图摘要（§3.6）；（6）智能体之间的对话由对话摘要（§3.7）促成。我们希望本文能阐明摘要技术在推动未来生成式智能体发展方面的潜力。

2 生成式智能体

生成式智能体是由人工智能驱动的计算软件，能够模拟可信的人类行为。本节简要概述生成式智能体架构及其几个主要组件。整体架构如图 1 所示。基础的**记忆**模块负责存储与智能体自身相关的各种信息，包括基本的观察以及高层的反思和生成的计划。随后，**检索**模块从记忆流中提取适当的记忆，以支持包括规划、反思和行动在内的下游模块。之后，**反思**模块为某个智能体的记忆提供高层抽象，这些抽象也构成另一种类型的记忆。此外，**规划**模块会考虑智能体的摘要和观测到的实体的摘要，并以一种精细的方式制定计划

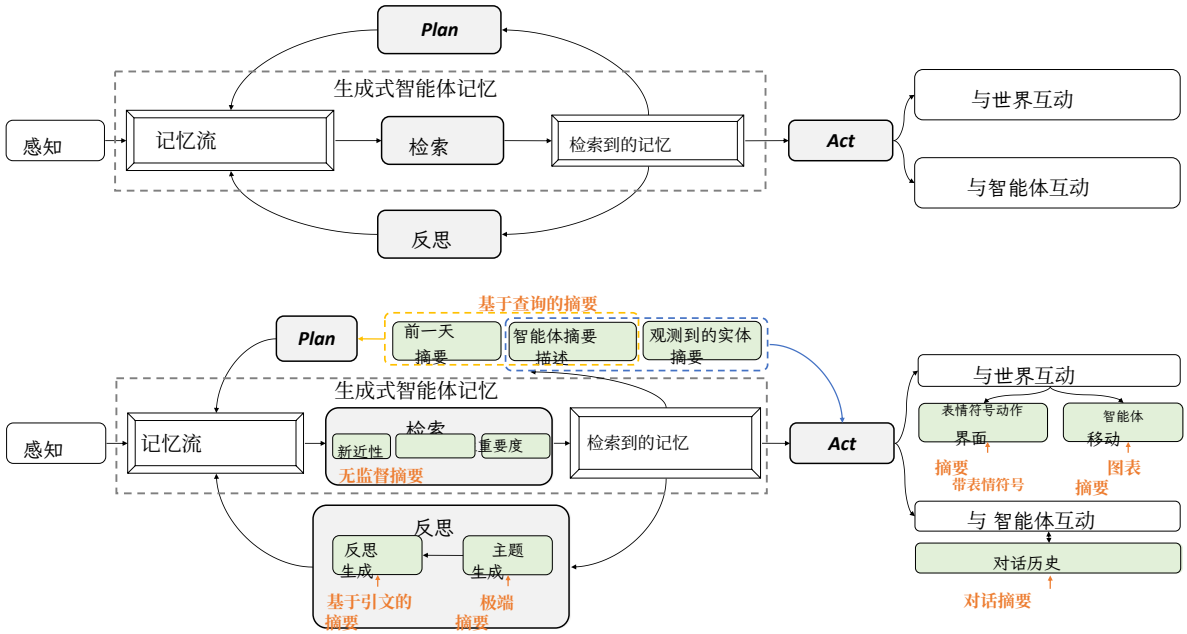


图1：插　图生成式智能体架构以及架构内部的关键摘要技术示意图　　e.

。最后，智能体通过执行动作与世界互动，或通过发起对话与其他智能体互动。

3 关键摘要技术

3.1 无监督摘要

检索模块旨在根据智能体当前的情况和整个记忆流，提供试点记忆。这与无监督摘要的目标一致：根据各种人工设计的特征，针对一个给定的查询，从一组文档中提取关键信息（Carbonell 和 Goldstein-Stewart，1998年）。

具体而言，检索函数会考虑三个不同的特征：新近性、重要性和相关性，以便从记忆流中有效提取重要信息。

具体来说，新近性 认为，最近访问过的记忆很重要，因为我们人类经常需要处理短期任务。重要性 由大语言模型根据其丰富的背景知识直接推断得出，具体方法是直接向大语言模型发出提示。相关性 会为与智能体当前情况最相关的记忆赋予更高的得分。最终的检索得分是三个得分的加权和：得分 $= \alpha_{\text{recency}} \cdot$ 新近性 $+ \alpha_{\text{importance}} \cdot$ 重要性 $+ \alpha_{\text{relevance}} \cdot$ 相关性。随着

上述三项特征的集成，检索模块成功对记忆进行无监督总结，为后续步骤生成摘要信息。

3.2 极端摘要

反思 是生成式智能体最关键的组成部分之一，它总结智能体近期的状况并形成高层次的思考。整个反思过程可以分为两个步骤，其中第一步是极端摘要。具体而言，该步骤旨在以问题生成的方式，将智能体最近的100条记忆记录浓缩为三个关键主题。

具体而言，该模块通过以下提示语引导大语言模型来实现这一目标：“仅根据上述信息，我们可以回答关于陈述中主体的哪3个最突出的高层次问题？”。结果包括三个高度浓缩的问题，这些问题可以被视为对智能体近期记忆的极端摘要，因为此前的多项研究证实了摘要与问题生成之间的紧密联系（Narayan 等人，2020；Feng 等人，2021）。

3.3 基于引文的摘要

第二步的**反思** 可以被视为一项基于引文的摘要任务，该任务接收若干带有索引的检索文档（记忆），旨在生成附有证据引用的摘要。这也与先前相关的工作生成任务以及开放式】【。

领域阅读理解任务，这两者都要求提供具体证据来支持其生成的结果（Chen 等人，2021）。

具体而言，该模块通过向大语言模型提示：“关于 Klaus Mueller 的陈述，……你能从上述陈述中推断出哪些高层次见解？（示例格式：见解（依据 1、5、3））”，来实现这一目标。输出将相关记忆概括为带有引文的反思：“Klaus Mueller 致力于研究绅士化（依据 1、2、8、15）”。

3.4 基于查询的摘要

事实上，在**检索** 模块的帮助下，基于查询的摘要贯穿整个生成式智能体的核心架构。在这一部分，我们主要关注三个任务，它们将为后续的**规划**和**行动** 模块提供支持。

智能体摘要描述 智能体摘要描述概括了智能体的身份信息、当前职业状况和自我评估，这些内容是制定计划和采取反应的重要线索。具体而言，首先通过三个查询“[name]的核心特征”、“[name]当前的日常职业活动”以及“[name’s] 对自己近期生活进展的感受”获取相关记忆，然后将生成的三个摘要合并，形成完整的智能体摘要描述。

前一日摘要 前一日摘要在计划制定过程中发挥着重要作用，可确保智能体实现连贯的长期目标。尽管原始论文中未提供详细信息（Park 等人，2023年），我们推测系统会使用隐式查询“[姓名]的前一日计划”来检索相关记忆并生成最终摘要。

被观察实体摘要 被观察实体摘要将（1）智能体与实体之间的关系以及（2）实体的状态进行压缩，是智能体是否采取动作的重要依据。该摘要由以下两个查询获得的部分组成：“ [观察者]与 [被观察实体]是什么关系？”以及“[被观察实体] 是 [被观察实体的动作状态]”。综合考虑智能体摘要描述和被观察实体摘要后，智能体通过向大语言模型提示“John 是否应对该观察作出反应？如果应该，什么样的反应才合适？”来决定是否作出反应。

3.5 表情符号摘要

为了快速了解智能体的状态，Park 等人（2023年）通过向大语言模型发送提示，在沙盒界面上实现了一种基于表情符号的高层抽象。例如，“Isabella Rodriguez 正在查看她的电子邮件”会显示为 。俗话说，一图胜千言，表情符号界面直观地概括了智能体当前的状态，并融入整个系统。

3.6 图摘要

生活在 Smallville 的智能体可以通过移动到达适当的位置。Smallville 采用树表示，其中根节点代表整个世界，子节点描述不同区域，叶节点则表示物体。智能体的移动方式首先通过将树表示转换为自然语言来确定，然后通过提示大语言模型回答“哪个区域应该[名称]前往？”。换言之，智能体的移动可以形式化为一项隐式图摘要任务（Kaushik，2003）。给定世界图，智能体会从当前位置出发，寻找一条通往目标目的地的合适路径。

3.7 对话摘要

智能体通过对话进行交互。在初始时刻，一个智能体会根据智能体摘要描述和被观察智能体（实体）的摘要，决定是否触发对话。为使对话连贯且内容丰富，后续话语会结合额外的对话摘要生成。在原始论文（Park 等人，2023年）中，纯对话历史被用于辅助生成对话。我们认为，面对冗长繁杂的对话历史时，对话摘要可以有效应对这一挑战。此外，在演示页面上，对话摘要也能让人快速了解对话的核心内容。

4 结论

在本报告中，我们旨在从统一的摘要视角理解生成式智能体。我们系统地分析了几种关键摘要技术，并展示了生成式智能体架构中的各个模块如何被形式化为传统摘要任务。我们认为，未来的生成式智能体可以通过先进的摘要能力得到显著提升。

参考文献

Jaime G. Carbonell 和 Jade Goldstein-Ste wart。1998年。使用 mmr（一种基于多样性的重新排序方法）对文档重新排序并生成摘要。载于年度国际 *ACM SIGIR* 信息检索研究与发展会议。

Xiuying Chen、Hind Alamro、Mingzhe Li、Shen Gao、Xiangliang Zhang、Dongyan Zhao 和 Rui Yan。2021 年。捕捉科学论文间的关系：用于生成相关工作部分的抽象式模型。载于第 *59* 届计算语言学协会年会暨第 *11* 届自然语言处理国际联合会议论文集（第 *1* 卷：长篇论文），第 6068–6077 页，线上举行。计算语言学协会。

Xiachong Feng、Xiaocheng Feng、Bing Qin 和 Xinwei Geng。2021 年。对话话语感知图模型与用于会议摘要的数据增强。载于第二十九届人工智能国际联合会议论文集（*IJCAI-21*），第 3808–3814 页。人工智能国际联合会议组织。主会场。

Shriraghav Kaushik. 2003. 用于索引图结构数据中路径的图摘要. 威斯康星大学麦迪逊分校。

Jinjie Mai, Jun Chen, Bing chuan Li, Guocheng Qian, Mohamed Elhoseiny, and Bernard Ghanem. 2023. 大语言模型作为机器人之脑：统一自我中心记忆与控制。*ArXiv*, abs/2304.09349。

Shashi Narayan, Gonçalo Simães, Ji Ma, Hannah Craighead, and Ryan T. McDonald. 2020. Qurious：面向文本生成的问题生成预训练。*ArXiv*，abs/2004.11026。

OpenAI. 2023. GPT-4 技术报告。*ArXiv*，abs/2303.08774。

Joon Sung Park, Joseph C O'Brien, Carrie J Cai, MeredithRingelMorris, Percy Liang, and Michael S Bernstein. 2023. 生成式智能体：人类行为的交互式模拟。*arXiv* 预印本 *arXiv:2304.03442*。

Chenfei Wu、Sheng-Kai Yin、Weizhen Qi、Xiaodong Wang、Zecheng Tang 和 Nan Duan。2023年。Visual ChatGPT：利用视觉基础模型进行对话、绘图和编辑。*ArXiv*, abs/2303.04671。

Jingfeng Yang、Hongye Jin、Ruixiang Tang、Xiaotian Han、Qizhang Feng、Haoming Jiang、Bing Yin 和 Xia Hu。2023年。实践中如何发挥大语言模型的力量：关于 ChatGPT 及更多内容的综述。

Wayne Xin Zhao、Kun Zhou、JunyiLi、Tianyi Tang、Xi- aolei Wang、Yupeng Hou、Yingqian Min、Beichen Zhang、Junjie Zhang、Zican Dong、Yifan Du、Chen Yang、Yushuo Chen、Z. Chen、Jinhao Jiang、Ruiyang Ren、Yifan Li、Xinyu Tang、Zikang Liu、Peiyu Liu、Jianyun Nie 和 Ji rong Wen。2023年。大型语言模型综述。*ArXiv*，abs/2303.18223。