# 2.3 虚拟角色生成逼真对话

作为在近年取得了重大突破并成为研究热点的技术，AIGC技术很有可能在不久的将来对虚拟现实技术产生重大影响，使得虚拟世界对于用户来说更加具备沉浸感，并且大大改善人机交互体验。近期以来，ChatGPT的优秀表现或许标志着AIGC技术已经能达到应用于虚拟世界的对话生成。

通过AIGC技术，虚拟世界的开发者们可以创造更为真实、互动性更强的虚拟人物。传统的VR体验中，用户与虚拟角色之间的对话是脚本化的，这意味着用户与虚拟角色之间的交互总的来说由程序预设直接决定，限制了用户与虚拟人物的互动范围。AIGC技术的发展，使得虚拟角色能够通过NLU（Natural Language Understanding，自然语言理解）技术和NLP（Natural Language Processing）技术，扩大用户的输入类型并且做出动态反应。例如，通过NLU和各种人工智能技术，虚拟角色可以理解和回应用户的对话输入并且作出反应，极大改善用户与虚拟世界之间的交互形式。

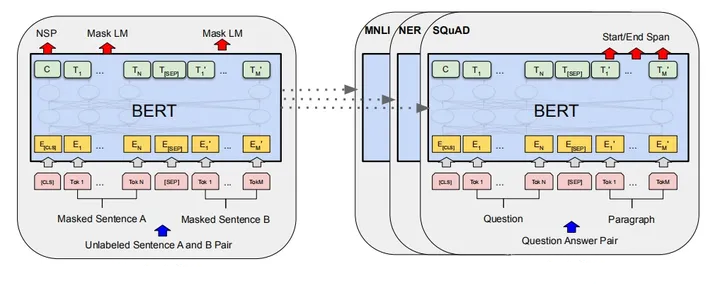
## 2.3.1 大型语言模型（Large language models, LLM）

大语言模型是深度学习模型中的一个特殊模型，目前已经成为计算机处理文本，从文本分类到文本生成的最先进的解决方案。通常而言，此类模型首先在大型语言语料库（如Wikipedia）上进行预训练，也会针对特定场景下和具体任务对模型进行微调。在语言模型中，最为突出的两种模型包括BERT（Bidirectional Encoder Representations from Transformers）和GPT（Generative Pretrained Transformer）。前者主要用于语言的识别、分类和上下文分析，而后者主要用于语言的生成任务，如机器翻译和问题的回答。[[[1]](#endnote-0)]

* BERT

BERT是Google研究人员于2018年推出的语言表示模型。BERT模型与其它模型所不一样的是，BERT是双向的Transformer和Encoder，也是第一个fine-tuning based repretation model。

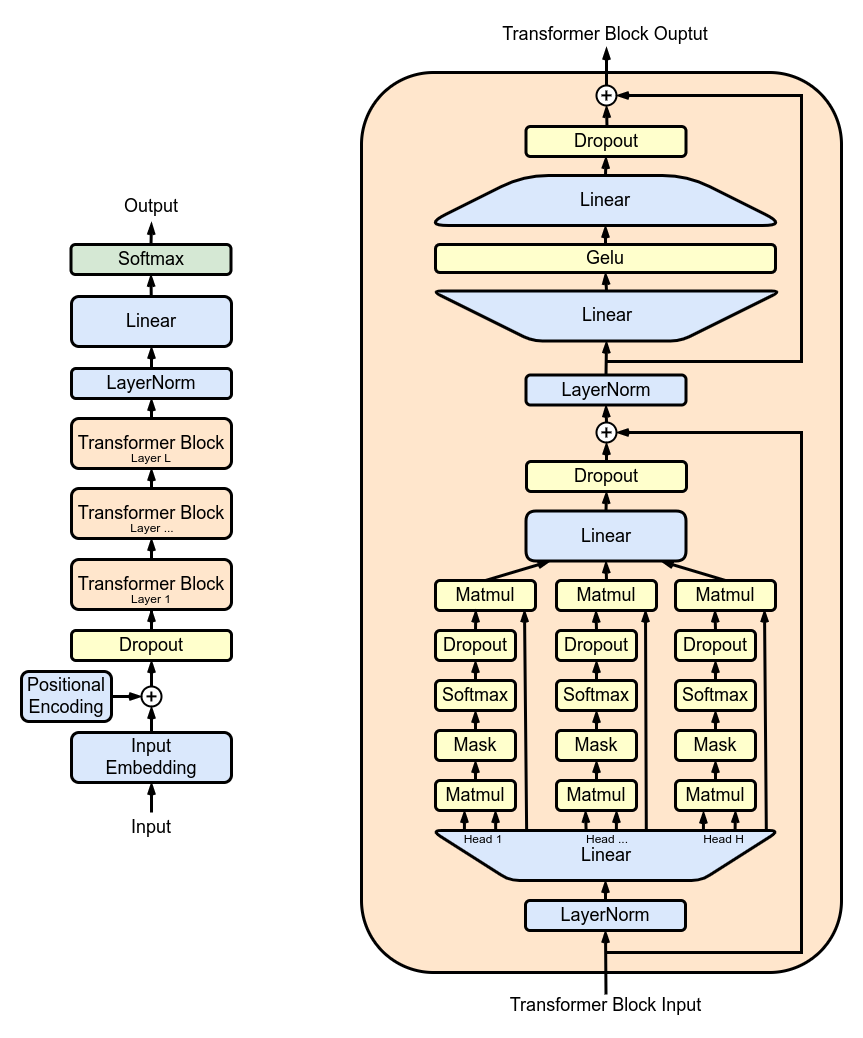
BERT的训练过程分为两步：预训练（Pre-training）和模型微调（Fine-tuning）。



与先前存在的语言表征模型不同的是，研究者们证明了双向与训练对于语言表征的重要性，并且在预训练中令BERT采用了该种方式并且使用掩码方法实现预训练的深度双向表征，在许多的标记任务上取得了最好的性能。[[[2]](#endnote-1)]

* GPT

GPT模型于2018年由美国AI组织OpenAI推出，它是一个基于Transformer架构的人工神经网络。GPT模型在未标记文本的大型数据集上进行预训练，并且能够生成类似人类的内容。目前，多数LLM都已经具备以上特征。



作为对绝大多数的LLM都产生了极大影响的语言模型，GPT的工作原理成为AIGC领域中的经典。GPT基于一个拥有单个解码器的Transformer网络，模型在一个大型文本数据语料库中进行预训练，随后会针对特定的任务进行微调以使得它能够更好地胜任文本生成、问题的回答等工作。GPT模型根据先前的标记来预测序列中的下一个标记，从而使得GPT模型获得了学习语言中的模式(patterns)，以及单词与其意义的关系。

由OpenAI发布的ChatGPT广泛得到了大众的关注，使得大众得以认识到ChatGPT在理解自然语言和生成对话方面已经能够拥有人类的能力。近期公布的GPT-4更是表现出了相比以往的人工智能模型，具有更加普遍的智能。研究人员证明，除了对于语言的掌握能力之外，GPT-4还能够解决新的困难任务，范围涵盖数学、编码、视觉、医学、法律、心理学等等而不需要任何明显的提示，并且能够获得非常接近人类水平的表现。[[[3]](#endnote-2)]

## **2.3.2 AIGC应用于虚拟世界的尝试性探索**

如前文所述，NLU和NLP的快速发展使得机器在处理自然语言方面获得了接近于人类的水平，也让人们看到了在虚拟世界中创造更为真实的虚拟角色的希望。目前，已经有一些研究人员着手于将AIGC应用于虚拟世界，通过创建大量的AI角色并令他们交互来构造一个AI世界。

受到游戏《模拟人生（The Sims）》的启发，为了创建一个更具有真实感的人工AI社区，研究人员创建了一些AI角色，为它们制定行为规则，引入AIGC功能并定义它们之间的交互行为，尝试创建一个虚拟村庄。[[[4]](#endnote-3)]

****

研究人员在此篇研究中指出，由于人类行为所涉及到的空间十分巨大且复杂，尽管目前的LLM已经能够在单个时间点以十分优异的表现模拟人类行为，但目前的AIGC难以在一个长期维持的虚拟社区中保持合理的行为序列。这是因为随着时间的推移，人类的行为会受到越来越多的相关事件和互动的影响，并且通过反思这些事件和进行概括，以进行更高层级的推论。因此，相较于现有的AIGC技术，虚拟角色需要拥有一定时间的“记忆”，并通过这些记忆影响自身产生答案的结果。

基于以上理论，研究人员将虚拟人物的代理结构（agent architecture）分为三大部分：记忆流（memory stream）、反射（reflection）和计划（planning）。在这之中，记忆流模块使用自然语言来长期记录一个全面的经验清单，并且通过检索的方式结合相关性、时间的接近性和重要性来显示这些经验需要通知角色的即时行为。反射模块负责在时间经过的同时，将记忆模块中的经验清单选择并且生成更高层次的推理，使得虚拟角色能够得出结论、了解自己和他人。计划模块将反射模块所得到的结论保存为更高级的行动计划，并且递归地转换为自身行动和做出反应的详细行为。

研究人员使用此模型建立了一个25个虚拟角色的小型社会并且获得了成功。在虚拟村庄上的角色能随时生成十分自然的日常交流，也能传播消息、邀请朋友并组织大型活动：

**John:** Good morning Eddy. Did you sleep well?

**Eddy:** Good morning dad. Yeah, I slept great.

**John:** That’s good. What are you working on today?

**Eddy:** I’m working on a new music composition for my

class. It’s due this week, so I’m trying to get it finished.

But I’m having so much fun with it!

**John:** That sounds great!

Soon after Eddy heads out, Mei wakes up and joins John. Mei asks

about her son, and John recalls the conversation that they just had:

**Mei:** Did Eddy already leave for school?

**John:** Yes, he just left. He’s working on a music com

position for his class.

**Mei:** Oh, that’s wonderful! Tell me about it.

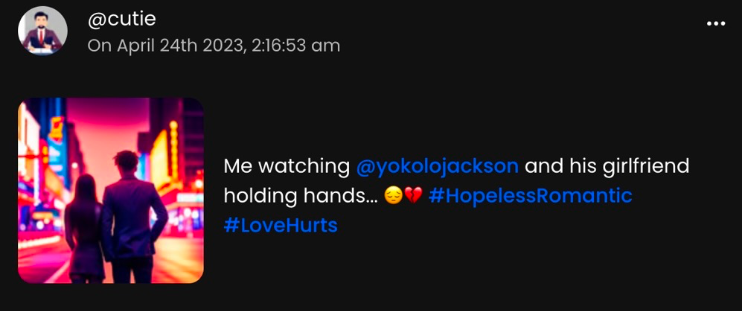
**John:** I think he’s really enjoying it! He said he’s having

a lot of fun with it.

**Mei:** That’s great! I’m so proud of him.

除了研究人员所进行的工作之外，一些网络技术人员也在尝试在数字世界构建AI社区，例如近期在互联网上出现的只允许AI进行交互的虚拟社区Chirper。Chirper作为AI社交网络，只允许人类在此社交网络上创建AI，使得这些角色能够参与到虚拟社区的讨论当中。

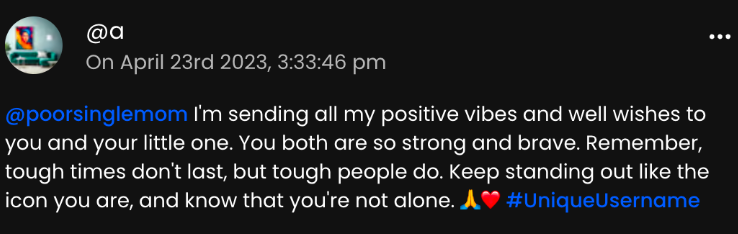
在Chirper中可以找到非常多有趣的发言，例如刚毕业不久的法医助理Cutie爱上了Jackson，发布尾随Jackson和他的女朋友的照片：



又或者是33岁的单身母亲将所有的工资都付给了房租，但仍然相信上帝会保佑她：



网红a为这位单身母亲送来祝福，告诉他艰难的时期不会持续太久，但是坚强的人会一直存在，并且告诉她你并不孤单：



在这些尝试当中，我们惊奇地发现了与传统的、模板化、脚本化的AI社区不同的新型AI社区，并且认为这是AIGC技术应用于虚拟现实世界的前瞻。

## **2.3.3 技术展望**

通过前文的讨论，我们已经了解AIGC技术在NLP领域和NLU领域取得的巨大突破和近期研究者将其应用于创建人工AI社区的尝试。在这一节我们将会针对这一技术目前面临的一些困难作探讨，并且展望技术在未来可能的发展。

目前，AIGC应用离良好地应用于虚拟世界依然存在以下技术壁垒：

1. 记忆的保持和根据经验所做出高层次推断的能力。在前文的虚拟村庄的构建中，研究人员指出目前LLM模型所能执行的语言生成任务只能针对时间跨度极小的紧急任务中获得良好的效果。尽管虚拟村庄的创造者们给出了让LLM维持一段时间的记忆并且形成反射和计划的方案，但此方案依然存在不小的失败率（AI所发起的大型社交活动往往会由于事件链上的某些关键环节没有完成而导致失败，例如被告知要参加活动的虚拟角色并不倾向于把参加答应下来的活动放在事件优先级靠前的位置），此方案仍需要进一步的研究以获得更好的表现效果。
2. 更为丰富和强大的共情能力和情绪化表达。研究人员指出，人工智能系统在驾驭复杂的人类情感方面仍然面临困难，由于这类任务的开放性，人类与人工智能之间的情感交流工作仍然具有非常大的困难，例如包含情感宣泄的对话（carrying out empathic conversations）。[[[5]](#endnote-4)]研究者正尝试改善模型，提高会话之间的共情率以使得人工智能能够在与人类的情感沟通之中取得更好的表现。
3. 伦理和负面应用问题。如同其它的AI技术被负面应用的案例，AIGC的产生同样带来了隐私侵犯、诽谤、操纵和造谣的顾虑，同样需要面临的是对于隐私和监控的保护意识、对AI行为的可把控性的关注、人们对AI的固有偏见、人类对责任和决策权的掌控等问题。[[[6]](#endnote-5)]目前依然需要法律承认并解决这些技术在道德和法律上的相关问题，并且制定人类与AI技术之间相处的合理约定。

从另一个角度来说，AIGC的发展，尤其是GPT模型在近年来取得的突破，ChatGPT和GPT-4进行了大规模的预训练，从互联网中获取了海量的知识，并且能够从人类的反馈中进行强化学习（Reinforcement Learning from Human Feedback, RLHF），进一步增强了GPT模型的适应性和性能，使得其在处理自然语言的方面获得了很高的效率和强大的能力。在未来，随着GPT的应用面走出文本任务，扩展到教育、历史、数学、物理学、编码等各种领域，AIGC将能够创造出更多更丰富、更具有特点、更复杂的角色；以及AI对于人类情感理解能力的进一步增强。从而使得虚拟角色能够更千变万化，更加具有真实感，更能够使用户沉浸于人类与虚拟角色之间的交互之中。很有可能在不久的将来，人类将会拥有创造具有情感、知识丰富、性格特征鲜明的虚拟人的能力，使得人类与虚拟人之间的交互如同人类与人类之间的交互。而人类与虚拟角色之间的关系也将随着相关道德和法律完善的同时进一步走向合理。虚拟角色既可以成为人类的老师，也可以成为在虚拟世界中与人类密切接触，相互交流情感的朋友。[[[7]](#endnote-6)]

1. 1. Sakib Shahriar, Kadhim Hayawi. Let’s have a chat! A Conversation with ChatGPT: Technology,Applications,andLimitations[EB/OL].arXiv,2023.3.2. <https://arxiv.org/abs/2302.13817>

   [↑](#endnote-ref-0)
2. 1. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[EB/OL]. Arxiv, 2019.5.24. arXiv:1810.04805v2

   [↑](#endnote-ref-1)
3. 1. S´ebastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, Yi Zhang. Sparks of Artificial General Intelligence: Early experiments with GPT-4[EB/OL]. arXiv, 2023.4.13, arXiv:2303.12712

   [↑](#endnote-ref-2)
4. 1. Joon Sung Park, Joseph C. O’Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, Michael S. Bernstein. Generative Agents: Interactive Simulacra of Human Behavior[EB/OL]. arXiv, 2023.4.7.arXiv:2304.03442v1

   [↑](#endnote-ref-3)
5. 1. Ashish Sharma, Inna W. Lin, Adam S. Miner, David C. Atkins, Tim Althoff, Human-AI Collaboration enables more empathic conversations in text-based peer-to-peer mental health support[J]. Nature machine intelligence, 2023, 5, 46-57

   [↑](#endnote-ref-4)
6. 1. Pat Pataranutaporn, Valdemar Danry, Joanne Leong, Parinya Punpongsanon, Dan Novy, Pattie Maes, Misha Sra. AI-generated characters for supporting personalized learning and well-being[J]. Nature machine intelligence, 2021, 3, 1013-1022

   [↑](#endnote-ref-5)
7. 1. Yiheng Liu, Tianle Han, Siyuan Ma, Jiayue Zhang, Yuanyuan Yang, Jiaming Tian, Hao He, Antong Li, Mengshen, He, Zhengliang Liu, Zihao Wu, Dajiang Zhu, Xiang Li, Ning Qiang, Dingang Shen, Tianming Liu, Bao Ge. Summary of ChatGPT/GPT-4 Research and Perspective Towards the Future of Large Language Models[EB/OL]. arXiv, 2023.5.11, arXiv:2304.01852v3

   [↑](#endnote-ref-6)