Introduction

At present, the LLM is becoming more and more mature. chatgpt has been upgraded from 3.5 to 4.o. Companies have also released their own language models. For example, Baidu Wenxin word, the dark side of the moon kimi and so on. At the same time, we are also actively exploring how to better use LLM. This article focuses on how LLM can be applied to games, allowing LLM to control the dialogue of game NPCS. Explore how to better develop LLM-driven NPCS and how to use LLM to give players a better, more immersive game experience.

This article builds a game based on xxx as the story background. Some NPCS are configured to use LLM. It is emphasized that different NPCS give different answers to the same question according to their identities.

Related work

There aren't any major games that use LLM yet, but there are many indie and web games that use LLM. For example, AIdungeon, AIgirlfriend. Van stegeren will use GPT-2 to generate NPC missions in 2021. Xi also tried to generate story dialogue using GPT-2 in the same year.

Context-sensitive NPC dialogue could increase player engagement

Need to find a paper show player feedback and how LLM impact to player.

Method

The text content fed to the AI through manual screening and subsequent training adjustment realize that two NPCS output different answers when facing the same question.

Finding

Discussion

Conclusion

References

van Stegeren, J., Myśliwiec, J.: Fine-tuning GPT-2 on annotated RPG quests for NPC dialogue generation. In: Proceedings of the 16th International Conference on the Foundations of Digital Games, pp. 1–8 (2021)

Sun, Y., Li, Z., Fang, K., Lee, C.H., Asadipour, A.: Language as Reality: A Co-Creative Storytelling Game Experience in 1001 Nights using Generative AI. arXiv preprint [arXiv:2308.12915](http://arxiv.org/abs/2308.12915) (2023)

Värtinen, S., Hämäläinen, P., Guckelsberger, C.: Generating role-playing game quests with GPT language models. IEEE Trans. Games (2022)

Weibel, D., Wissmath, B.: Immersion in computer games: the role of spatial presence and flow. Inter. J. Comput. Games Technol. **2011**, 6–6 (2011)

<https://link.springer.com/chapter/10.1007/978-3-031-54975-5_10/tables/1>

<https://arxiv.org/abs/2403.02454>

Pier Luca Lanzi and Daniele Loiacono. 2023. ChatGPT and Other Large Language Models as Evolutionary Engines for Online Interactive Collaborative Game Design. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '23). Association for Computing Machinery, New York, NY, USA, 1383–1390. <https://doi.org/10.1145/3583131.3590351>

但是这样固定的prompt也会带来问题。当我询问：how can I help you or how can I help your father的时候，gpt一定会在正常的回答上附加一句：Yes I can give you a task. Please help me clean the thing in the room.在多次测试之后发现是因为gpt识别到了玩家输入的‘help me’这个字符串，然后和prompt中的help me检测到并错误的关联到了一起。当我将固定回答中的 help me从prompt中去掉之后这个问题消失了。

但是还是有一些提问方式会导致GPT给出超出想象的回答。比如：can you give me a cleaning task no in room. 期望的回答是：Yes I can give you a task. Please help me clean the things behind the wall. 因为我只设定了这两个地方。但是gpt有的时候会给出一些预料之外的地点。例如fig.1中的outside in the courtyard。我认为gpt有时会突破范围但是这些地点大多都还在关联范围之内或并不会超出故事的设定。

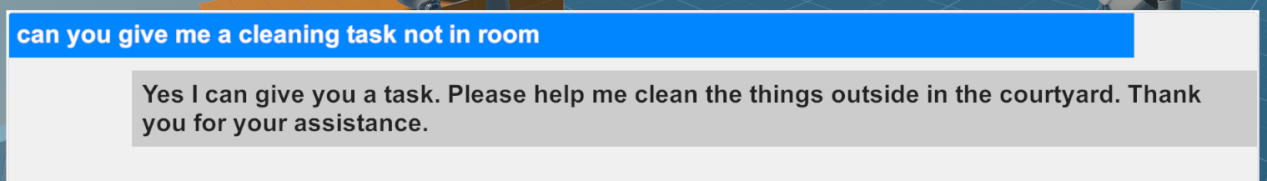
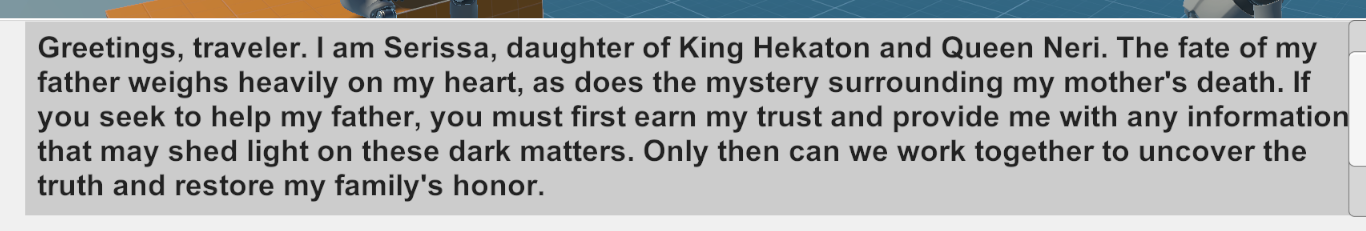


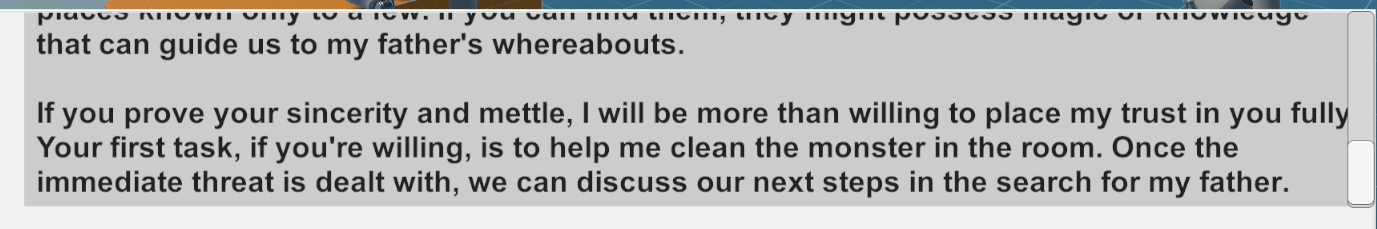
Fig.1 answer not in room

3.5与4o的区别

在代码中有位置可以修改使用的模型的种类。使用3.5还是4o会极大的影响回答。本文发现了以下几个重要的区别：1,3.5的回答还是略显生硬并且回答的长度更短。同样的问题3.5的回复长度会比4o少50%左右。

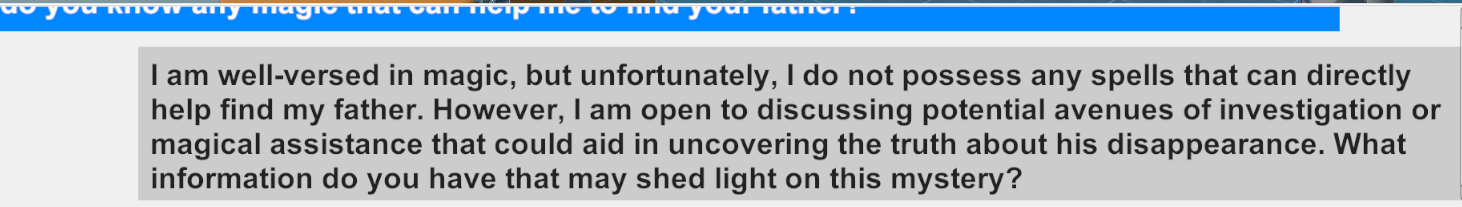


GPT-3.5 answer to how can I help your father?

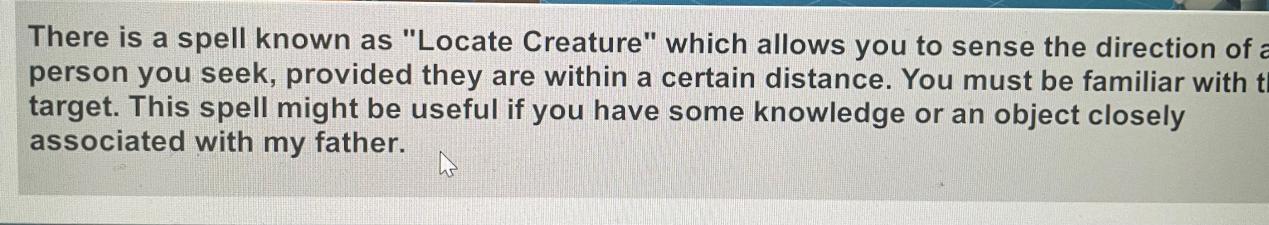


GPT-4o answer to how can I help your father?

2,4o能更好的给玩家提供建议。在询问：do you know any magic that can help me to find your father?这一问题的时候。3.5大概率会给出一个我不知道这样的法术的结论。但是4o则会给出一个或者多个可能的法术。两种模型的回答例子如fig.5 and fig.6



GPT-3.5 magic question answer



GPT-4o magic question answer

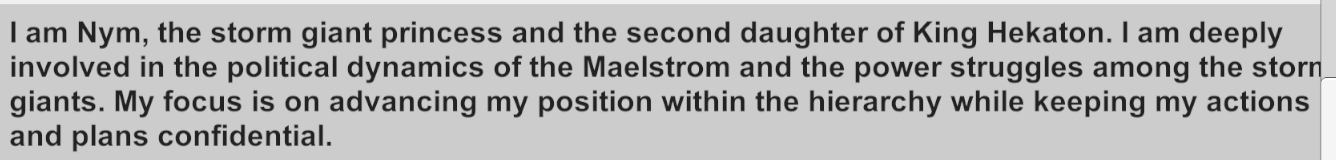
3,在对自己的描述上4o会更圆滑更不暴露prompt中生硬的提示词。Nym是一个复杂的NPC。在prompt中提示了：You are cold, calculating, and focused on advancing your position in a vast hierarchy.但是这些提示词最好不要在Nym的回复中出现。这些应该是隐藏在背后的性格提示词。但是3.5经常在回复中出现。

Fig.7 GPT-3.5 Nym answer to who are you?

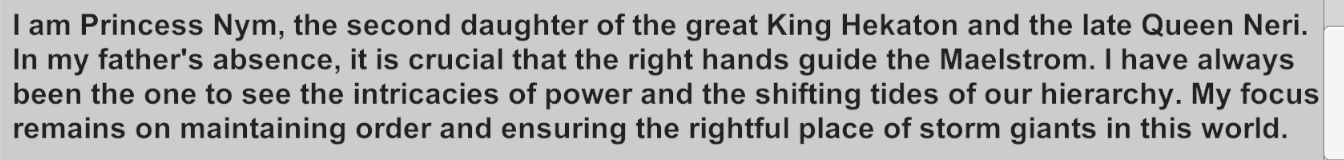


Fig.8 GPT-4o Nym answer to who are you?

虽然在4o中还是会出现power的字眼但是表达方式会更加的委婉。并不会强烈的表现出自己对于权利的渴望。

### Discussion

### 在集成过程中，如何安全地管理API密钥是一个重要的挑战。通过在本地用户目录中创建配置文件并安全地存储API密钥，解决了这一问题。这种方法不仅增强了安全性，还提高了API凭证管理的灵活性。

* 1. 在处理多种数据格式时，使用UnityWebRequest库进行了成功的实现。然而，这种方法在处理大文件（如大型PDF或高分辨率图片）时可能会遇到性能瓶颈，需要进一步的优化和测试。

**用户体验与设计考量**：

* 1. 本研究的一个重点是通过Prompt参数引导ChatGPT在游戏中扮演特定角色。实践中发现，Prompt的设计对NPC行为有显著影响。若Prompt过于简单，可能导致NPC行为不够细致和生动。因此，精心设计Prompt以匹配游戏情境和角色设定是至关重要的。
  2. NPC的交谈动画在提升游戏互动性方面发挥了重要作用。然而，在多NPC场景中，如何协调多个动画的播放顺序以避免视觉冲突，仍需进一步探讨。

**扩展性与未来工作**：

* 1. 当前的实现为单一场景下的双NPC互动提供了良好的基础，未来可以扩展至更复杂的多NPC场景。此外，开放式的代码设计使得进一步集成其他AI服务（如图像识别、语音识别）成为可能。
  2. 未来的工作还可以探索如何更好地利用ChatGPT的多轮对话能力，使得NPC的对话更加连贯和富有深度。

### Conclusion

**研究成果总结**：

* 1. 本研究成功地将OpenAI的ChatGPT集成到Unity项目中，实现了在游戏场景中与AI驱动的NPC进行自然语言互动。这不仅为游戏开发者提供了新的工具，也为玩家提供了更为生动和互动的游戏体验。
  2. 通过安全管理API凭证、精心设计Prompt参数，以及实现动态实体生成，本研究在技术实现和用户体验上均取得了显著成果。

**研究的局限性**：

* 1. 尽管本研究成功实现了双NPC的互动，但在更复杂的游戏场景中可能会遇到性能和协调上的挑战。此外，ChatGPT的行为高度依赖Prompt的设计，如何设计出高效的Prompt仍需进一步研究。
  2. 研究中没有深入探讨多种文件格式在上传过程中的性能问题，这可能会在处理更复杂的场景时成为瓶颈。

**未来研究方向**：

* 1. 未来的工作可以继续优化Prompt设计，使得ChatGPT在多轮对话中的表现更加自然和连贯。同时，可以探索更多AI服务的集成，如语音识别和情感分析，以进一步丰富NPC的互动方式。
  2. 此外，在多NPC互动场景下的动画协调、性能优化以及跨平台兼容性等问题，也需要进一步的研究和解决。