

MINISTRY OF EDUCATION AND RESEARCH OF THE REPUBLIC OF MOLDOVA

Technical University of Moldova

Faculty of Computers, Informatics and Microelectronics

Department of Software and Automation Engineering

Admitted to defense

Department Head:

FIODOROV Ion dr., conf.univ.

„___” _____ 2024

**THE INTEGRATION OF LARGE LANGUAGE
MODELS INTO INTELLIGENT AGENTS FOR
ENHANCED INTERACTIVE STORYTELLING**

Master's thesis

Student: _____ **Vrabie Teodor, IS-221M**

Coordinator: _____ **Braga Vasili, lect. univ.**

Consultant: _____ **Catruc Mariana, lect.univ.**

Chişinău, 2024

ADNOTARE

Această teză explorează crearea și implementarea unui sistem interactiv de povestire prin integrarea perfectă a Unity, API-ului OpenAI și Unity Machine Learning Agents (ML-Agents). Obiectivul principal este furnizarea unei experiențe captivante și immersive de povestire care se adaptează dinamic la interacțiunile utilizatorului. Unity, un motor puternic de dezvoltare a jocurilor, servește ca element de bază, oferind un mediu pentru lumi vizual atractive și elemente dinamice ale poveștii. Integrarea API-ului OpenAI aduce capacități avansate de prelucrare a limbajului natural, deblocând potențialul generării de text asemănător cu cel uman și facilitând povestirea relevantă contextual. Visual Studio, ca mediul integrat de dezvoltare (IDE), optimizează procesul de dezvoltare, oferind instrumente de depanare și capacități de colaborare.

Includerea Unity ML-Agents introduce inteligența artificială în sistem, permițând personajelor și entităților din lumea virtuală să învețe, să se adapteze și să răspundă interacțiunilor utilizatorului în timp real. Această inteligență adaptivă îmbunătățește calitatea imersivă a narativei, creând o experiență dinamică și personalizată pentru fiecare utilizator. Realizarea proiectului implică depășirea provocărilor asociate cu Unity, cum ar fi o curbă de învățare abruptă, probleme de performanță și inconsistente în documentație. În plus, se abordează dependențele de platformă, calitatea magazinului de active, limitele grafice, costurile de licențiere, intensitatea resurselor și provocările colaborării.

Detaliile de implementare cuprind traducerea proprietăților de semnalizare din lumea reală în semne virtuale cu diferențe în lizibilitate și dimensiunea textului. Personajele din mit sunt reprezentate grafic, iar camere virtuale sunt atașate pentru o vedere în prima persoană. Dinamica navigației în lumea mitică implică verificări de vizibilitate cu raze încântătoare care influențează revelarea semnificației semnului. Calculările de vizibilitate evaluează esența semnului, luând în considerare atât vizibilitatea, cât și factorii de lizibilitate. Teza se încheie cu o convergență a creativității, tehnologiei și inteligenței artificiale, formând baza pentru o experiență interactivă de povestire în care alegerile utilizatorului dau forma narativei în moduri fără precedent. Această teză nu abordează doar aspectele tehnice ale integrării Unity, OpenAI și ML-Agents, ci se adâncește și în aspectele creative ale povestirii într-un mediu virtual. Prin depășirea provocărilor asociate cu dezvoltarea de jocuri și valorificarea tehnologiilor de vârf, proiectul își propune să redefinească implicarea utilizatorului în cadrul poveștilor interactive. Fusionarea prelucrării avansate a limbajului natural, a inteligenței artificiale adaptive și a unui motor solid de dezvoltare a jocurilor promite să transforme paradigmele tradiționale de povestire. Prin implementare meticuloasă și abordări inovatoare ale elementelor interactive, această cercetare contribuie la peisajul în evoluție al experiențelor ghidate de narativă, punând bazele pentru progrese viitoare în povestirea digitală.

ANNOTATION

This thesis explores the creation and implementation of an interactive storytelling system by seamlessly integrating Unity, the OpenAI API, and Unity Machine Learning Agents (ML-Agents). The core objective is to deliver a captivating and immersive narrative experience that adapts dynamically to user interactions. Unity, a powerful game development engine, serves as the foundational element, providing an environment for visually engaging worlds and dynamic story elements. The integration of the OpenAI API introduces advanced natural language processing capabilities, unlocking the potential for generating human-like text and enabling contextually relevant storytelling. Visual Studio, as the Integrated Development Environment (IDE), streamlines the development process, offering debugging tools and collaborative capabilities.

The inclusion of Unity ML-Agents introduces artificial intelligence to the system, allowing characters and entities within the virtual world to learn, adapt, and respond to user interactions in real-time. This adaptive intelligence enhances the immersive quality of the narrative, creating a dynamic and personalized experience for each user. The project's realization involves overcoming challenges associated with Unity, such as a steep learning curve, performance issues, and documentation inconsistencies. Additionally, platform dependencies, asset store quality, graphics limitations, licensing costs, resource intensiveness, and collaboration challenges are addressed.

Implementation details encompass the translation of real-world signage properties into virtual signs with differences in readability and text size. Characters from the myth are graphically represented, and virtual cameras are attached for a first-person view. Navigation dynamics in the mythical realm involve visibility checks with enchanted rays influencing the revelation of sign significance. Visibility calculations gauge the essence of the sign, considering both visibility and readability factors. The thesis concludes with a convergence of creativity, technology, and artificial intelligence, forming the foundation for an interactive storytelling experience where user choices shape the unfolding narrative in unprecedented ways. This thesis not only addresses the technical intricacies of integrating Unity, OpenAI, and ML-Agents but also delves into the creative aspects of storytelling in a virtual environment. By overcoming challenges associated with game development and leveraging cutting-edge technologies, the project aims to redefine user engagement within interactive narratives. The fusion of advanced natural language processing, adaptive artificial intelligence, and a robust game development engine holds promise for transforming traditional storytelling paradigms.

Through meticulous implementation and innovative approaches to interactive elements, this research contributes to the evolving landscape of narrative-driven experiences, laying the groundwork for future advancements in immersive digital storytelling.

Contents

INTRODUCTION	8
1. ANALYSIS OF THE STUDY DOMAIN	9
1.1 Overview of Large Language Models.....	10
1.2 Navigating Competitor Territory: A Strategic Analysis	11
1.3 Exploring the Framework and Components of Large-Scale Language Models	14
1.4 Training Pipeline: A Quick Overview	17
1.5 Challenges and Limitations	24
2. BUILDING BLOCKS: PROJECT LANGUAGES AND TECHNOLOGIES	27
2.1 ChatGPT: Enhancing Storytelling and Interaction	28
2.2 Crafting Engaging Narratives in Collaborative Storytelling.....	29
2.3 Ensuring Model Excellence: Comprehensive Testing Approaches for Language Models....	35
2.4 Challenges in Collaborative Narratives with Large-Scale Neural Language Models	39
3. REALIZATION OF THE SYSTEM	42
3.1 Unity.....	42
3.2 Implementation	44
3.3 Challenges and limitations	48
3.4 Testing the system.....	50
CONCLUSION.....	53
BIBLIOGRAPHY	54

INTRODUCTION

Storytelling holds a central role in human socialization and entertainment, with various traditional forms like novels, plays, television, and movies offering passive audience experiences throughout history. However, gaming introduces a unique dynamic due to its emphasis on interactivity. Unlike other mediums, games involve the audience actively, presenting a challenge in finding the right balance between player freedom and a well-defined storyline. While excessive player freedom might lead to unexplored narratives, too many restrictions risk turning gaming into a passive experience. Striking this delicate balance is a significant design challenge in gaming, requiring careful consideration to ensure both engaging gameplay and a compelling narrative.

As gaming technology progresses, new prospects emerge for interactive storytelling. Improved storage capabilities enable the narration of longer, more intricate stories, while enhanced graphical features contribute to more immersive gaming encounters. Progress in artificial intelligence results in more formidable opponents, lifelike non-player character behavior, and other advancements. Superior procedural content generation algorithms ensure unique and enduring gameplay experiences. Notably, recent breakthroughs in language modeling offer a novel opportunity: the potential for on-demand generation of language and stories.

This paper introduces an innovative collaborative storytelling game, where a human player and an artificial intelligence agent collaboratively construct a narrative. The game begins with the AI agent presenting one of several curated story starters—opening sentences designed to spark participants' creativity. The human player responds by contributing a line, termed a story continuation, and the AI agent and human player alternately add continuations until the human player concludes the story. This game is intentionally designed with minimal restrictions, differing from traditional storytelling settings where the narrative is predetermined.

Our key contributions can be summarized as follows:

- We propose a unique endeavor in collaborative storytelling, where the collaboration between humans and AI agents is leveraged to craft a narrative collectively.
- We introduce a collaborative storytelling system that is crafted by fine-tuning a large-scale neural language model using a dataset of writing prompts stories.
- We devise a methodology for evaluating and ranking the language model's outputs, aiming to generate story continuations that closely resemble human-like expressions.
- We perform a comprehensive analysis, both quantitative and qualitative, of our system's storytelling capabilities by engaging in collaborative storytelling sessions with humans

BIBLIOGRAPHY

1. C. Tillmann and H. Ney, “Word Reordering and a Dynamic Programming Beam Search Algorithm for Statistical Machine Translation,” *Computational Linguistics*, vol. 29, no. 1, pp. 97–133, Mar. 2003, doi: 10.1162/089120103321337458.
2. Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin, “A Neural Probabilistic Language Model”.
3. A. Vaswani et al., “Attention is All you Need”. Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. *Language Models are Unsupervised Multitask Learners*. (2019).
4. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” *arXiv*, May 24, 2019. Accessed: Jan. 09, 2024. [Online]. Available: <http://arxiv.org/abs/1810.04805>
5. L. Yao, N. Peng, R. Weischedel, K. Knight, D. Zhao, and R. Yan, “Plan-and-Write: Towards Better Automatic Storytelling,” *AAAI*, vol. 33, no. 01, pp. 7378–7385, Jul. 2019, doi: 10.1609/aaai.v33i01.33017378.
6. P. Gupta, V. Bannihatti Kumar, M. Bhutani, and A. W. Black, “WriterForcing: Generating more interesting story endings,” in *Proceedings of the Second Workshop on Storytelling*, Florence, Italy: Association for Computational Linguistics, 2019, pp. 117–126. doi: 10.18653/v1/W19-3413.
7. F. Luo et al., “Learning to Control the Fine-grained Sentiment for Story Ending Generation,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy: Association for Computational Linguistics, 2019, pp. 6020–6026. doi: 10.18653/v1/P19-1603.
8. “Information Sciences Institute University of Southern California,” *AI Magazine*, vol. 1, no. 1, pp. 22–25, Jan. 1980, doi: 10.1609/aimag.v1i1.88.
9. Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin, “A Neural Probabilistic Language Model”.
10. Real-time Animated Storytelling - Unity Learn. <https://learn.unity.com/course/real-time-animated-storytelling>
11. A. Happe and J. Cito, “Getting pwn’d by AI: Penetration Testing with Large Language Models,” in *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, San Francisco CA USA: ACM, Nov. 2023, pp. 2082–2086. doi: 10.1145/3611643.3613083.
12. J. Kocoń et al., “ChatGPT: Jack of all trades, master of none,” *Information Fusion*, vol. 99, p. 101861, Nov. 2023, doi: 10.1016/j.inffus.2023.101861.
13. J. Wang et al., “Review of large vision models and visual prompt engineering,” *Meta-Radiology*, vol. 1, no. 3, p. 100047, Nov. 2023, doi: 10.1016/j.metrad.2023.100047.
14. H. J. Simon and E. J. Yin, “Microbioassay of Antimicrobial Agents,” *APPL. MICROBIOL.*, vol. 19, 1970.

15. D. L. Roberts, M. O. Riedl, and C. L. Isbell, “Beyond Adversarial: The Case for Game AI as Storytelling”.
16. M. L. Tielman, M. A. Neerinx, R. Bidarra, B. Kybartas, and W.-P. Brinkman, “A Therapy System for Post-Traumatic Stress Disorder Using a Virtual Agent and Virtual Storytelling to Reconstruct Traumatic Memories,” *J Med Syst*, vol. 41, no. 8, p. 125, Aug. 2017, doi: 10.1007/s10916-017-0771-y.
17. D. L. Martin, A. J. Cheyer, and D. B. Moran, “The open agent architecture: A framework for building distributed software systems,” *Applied Artificial Intelligence*, vol. 13, no. 1–2, pp. 91–128, Jan. 1999, doi: 10.1080/088395199117504.
18. C. Easter, “Virtual Environments for Research into Social Evolution (VERSE): A novel experimental environment for the study of human social learning,” *Animal Behavior and Cognition*, preprint, Aug. 2022. doi: 10.1101/2022.08.24.505099.
19. D. Thue, V. Bulitko, M. Spetch, and E. Wasylishen, “Interactive Storytelling: A Player Modelling Approach,” *AIIDE*, vol. 3, no. 1, pp. 43–48, Sep. 2021, doi: 10.1609/aiide.v3i1.18780.
20. S. Kim, S. Moon, S. Han, and J. Chang, “Programming the Story: Interactive Storytelling System”