MINISTRY OF EDUCATION AND RESEARCH OF THE REPUBLIC OF MOLDOVA
Technical University of Moldova
Faculty of Computers, Informatics and Microelectronics
Department of Software Engineering and Automatics

Admitted to defense
Department chief:
FIODOROV Ion dr., conf.univ.

„___” _____ 2024

THE INTEGRATION OF THEORY OF MIND INTO INTELLIGENT AGENTS FOR ENHANCED INTERACTIVE STORYTELLING

Master's project

Student: _____ **Peticî Marina, IS-221M**

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

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

Chişinău, 2024

ADNOTARE

Această teză are drept scop cercetarea posibilităților de integrare a Teoriei Minții în agenții inteligenți pentru îmbunătățirea povestirilor interactive digitale. Lucrarea evidențiază modalități prin care agenții inteligenți pot fi echipați cu caracteristici umane, permițându-le să înțeleagă și să empatizeze mai eficient atât cu utilizatorii, cât și cu alți agenți din mediul virtual.

Capitolul întâi stabilește fundamentul teoretic al lucrării, introducând conceptul de Teorie a Minții și importanța sa în povestirea interactivă. Se analizează actualitatea dezvoltării în narațiunea interactivă și Teoria Minții, evidențiind provocările existente și posibilele soluții.

Capitolul doi descrie în detaliu specificațiile sistemului propus pentru implementare. Capitolul se axează pe arhetipurile narative: Eroul, Mentorul și Răufăcătorul și modul în care Teorii Minții poate îmbunătăți comportamentul agenților inteligenți în cadrul poveștilor virtuale.

Capitolul trei detaliază implementarea practică a sistemului în mediul Unity, punând accent pe dezvoltarea interfeței utilizator și a mecanicilor de interacțiune. Este descris despre modul în care utilizatorii interacționează cu lumea narativă și personajele sale, și se prezintă o simulare a dialogului între personaje pentru a demonstra aplicarea Teoriei Minții în practică.

Cercetarea arată că integrarea Teoriei Minții în agenții inteligenți conduce la interacțiuni mai naturale și mai relevante din punct de vedere emoțional. Sistemul dezvoltat permite agenților inteligenți să interpreteze și să răspundă adecvat la stările emoționale și intențiile utilizatorilor. Agenții inteligenți, modelați după arhetipurile narative tradiționale, sunt îmbunătățiți cu reguli din Teoria Minții, reflectând astfel o înțelegere profundă a contextului narativ și a dinamicilor interpersonale.

Sistemul este construit utilizând Unity, un motor de joc puternic și versatil, care a permis implementarea avansată a interfețelor utilizator și a mecanicilor de joc. A fost folosită tehnologia de procesare a limbajului natural, încorporând API-ul ChatGPT pentru facilitarea interacțiunilor bazate pe text între utilizatori și agenții inteligenți. Este inclus un set de scripturi care gestionează mișcarea personajelor, interacțiunile și răspunsurile, asigurând o experiență de povestire fluidă și imersivă.

În concluzie, teza demonstrează modul în care integrarea avansată a Teorii Minții și a Inteligenței Artificiale în povestirea interactivă poate revoluționa modul în care poveștile sunt spuse și experimentate în medii digitale. Aceasta deschide noi orizonturi pentru dezvoltarea jocurilor și a aplicațiilor educaționale, oferind o experiență narativă mai profundă și mai imersivă pentru utilizatori.

ANNOTATION

This thesis aims to explore the possibilities of integrating Theory of Mind into intelligent agents to enhance digital interactive storytelling. The work highlights ways in which intelligent agents can be equipped with human-like characteristics, allowing them to understand and empathize more effectively with both users and other agents in the virtual environment.

The first chapter establishes the theoretical foundation of the work, introducing the concept of Theory of Mind and its importance in interactive storytelling. It examines the current state of the art in interactive narration and Theory of Mind, highlighting existing challenges and possible solutions.

The second chapter describes in detail the specifications of the proposed system for implementation. The chapter focuses on narrative archetypes: the Hero, the Mentor, and the Villain, and how Theory of Mind can improve the behavior of intelligent agents within virtual stories.

The third chapter details the practical implementation of the system in the Unity environment, emphasizing the development of the user interface and interaction mechanics. It describes how users interact with the narrative world and its characters and presents a simulation of dialogue between characters to demonstrate the application of Theory of Mind in practice.

The research shows that integrating Theory of Mind into intelligent agents leads to more natural and emotionally relevant interactions. The developed system allows intelligent agents to interpret and respond appropriately to users' emotional states and intentions. Intelligent agents, modeled after traditional narrative archetypes, are enhanced with rules from Theory of Mind, thereby reflecting a deep understanding of the narrative context and interpersonal dynamics.

The system is built using Unity, a powerful and versatile game engine, which allows for the advanced implementation of user interfaces and game mechanics. Natural language processing technology, incorporating the ChatGPT API, is used to facilitate text-based interactions between users and intelligent agents. The system includes a set of scripts that manage character movement, interactions, and responses, ensuring a fluid and immersive storytelling experience.

In conclusion, the thesis demonstrates how the advanced integration of Theory of Mind and Artificial Intelligence in interactive storytelling can revolutionize the way stories are told and experienced in digital media. This opens new horizons for the development of games and educational applications, offering a more profound and immersive narrative experience for users.

TABLE OF CONTENTS

INTRODUCTION

In the age of rapid technological evolution, the boundaries of artificial intelligence (AI) and human interaction are consistently being redrawn. The internet has revolutionized many aspects of life, including the age-old practice of storytelling which has transitioned from the traditional oral tales shared around a campfire to engaging digital narratives experienced on screens. At the core of this evolution is the goal to evoke emotions, foster connections, and deepen understanding among listeners. Nowadays, storytelling goes beyond merely narrating events because it seeks to actively involve the audience in the narrative. The instrument to achieve this is an Interactive Storytelling Platform, which exemplifies the blend of technology and storytelling, transforming users from readers into active participants in the unfolding story. This constantly changing format of storytelling not only reshapes how stories are told but also how they are perceived and interacted with by the audience. The depth of engagement offered by interactivity is not just a technological advancement, but a fundamental alteration of the psychological dynamics of storytelling. Here, the concept of Theory of Mind (ToM) becomes particularly relevant. The Theory of Mind, a concept deeply rooted in cognitive psychology, refers to the ability of an individual to attribute mental states to oneself and others. It's the understanding that others have beliefs, aspirations, emotions, and intentions that might be different from one's own [1].

For digital narratives, this psychological aspect is indispensable. As users engage with the stories, they are not just passive participants, but are actively interpreting and predicting the mental states of characters, delving into complex emotional and social scenarios. This process mirrors the cognitive functions involved in Theory of Mind, where understanding and empathy are developed through the exploration of diverse perspectives.

In the case of human-human interactions, ToM enables empathy, deception and cooperation. However, when it comes to human-AI interactions, especially in the context of interactive storytelling, the technology available nowadays cannot genuinely comprehend user's emotions and dynamically respond, in a feedback chain, to user actions. This leads to the central questions of this thesis: Is it possible to further enhance human-AI interactions? Can an AI-driven character be developed to adapt the progression of a story based on the user's perceived mood or intentions? Additionally, can it understand the thoughts, beliefs, and desires of other AI-driven characters and real users?

This thesis aims to explore the intersection of interactive storytelling, artificial intelligence, and the Theory of Mind, particularly focusing on how they function together within simulation games. The objective is to achieve more credible responses from intelligent agents during interactions in a virtual environment. Through this exploration, the ultimate goal is to understand the possibilities and challenges involved in integrating digital narratives with a deeper sense of a humane understanding. This research draws inspiration from basic storytelling archetypes and the principles of Theory of Mind.

REFERENCES

[1] SCIENCEDIRECT: *Theory of Mind*. Science Direct, @2023 [online]. Available: https://www.sciencedirect.com/science/article/pii/S187892931100048X

[2] Colombo M, Piccinini G. The Computational Theory of Mind. Cambridge: Cambridge University Press; 2023. doi:10.1017/9781009183734 [online]. Available: https://www.structural-learning.com/post/theory-of-mind

[3] Rescorla, Michael, "The Computational Theory of Mind", The Stanford Encyclopedia of Philosophy (Fall 2020 Edition), Edward N. Zalta (ed.). [Online]. Available: https://plato.stanford.edu/entries/computational-mind/?ref=jack-chong

[4 Copeland, J., 1996, "What is Computation?", Synthese, 108: 335–359

[5] Stevens, A. (2012). The archetypes. The handbook of Jungian psychology (pp. 74-93) [online]. Available: https://www.routledge.com/The-Handbook-of-Jungian-Psychology-Theory-Practice-and-Applications/Papadopoulos/p/book/9781583911488

[6] Petrovic M, Bonanno S, Landoni M, Ionio C, Hagedoorn M, Gaggioli A. Using the Transformative Storytelling Technique to Generate Empowering Narratives for Informal Caregivers: Semistructured Interviews, Thematic Analysis, and Method Demonstration. JMIR Form Res. 2022 Aug 2;6(8):e36405. doi: 10.2196/36405. Erratum in: JMIR Form Res. 2022 Sep 14;6(9):e42323. PMID: 35802492; PMCID: PMC9382549.

[7] TWINERY: *Welcome to the Twine Cookbook*. Twinery.org, @2021 [accessed 14.09.2023]. Available: https://twinery.org/cookbook/

[8] Green, C.P., Holmquist, L.E., Gibson, S. (2020). Towards the Emergent Theatre: A Novel Approach for Creating Live Emergent Narratives Using Finite State Machines. In: Bosser, AG., Millard, D.E., Hargood, C. (eds) Interactive Storytelling. ICIDS 2020. Lecture Notes in Computer Science(), vol 12497. Springer, Cham. Available: https://www.researchgate.net/figure/Story-structure-in-Twine_fig2_337706357

[9] Towards the Emergent Theatre: A Novel Approach for Creating Live Emergent Narratives Using Finite State Machines. Available: https://link.springer.com/chapter/10.1007/978-3-030-62516-0_8

[10] https://www.tutorialspoint.com/digital_circuits/digital_circuits_finite_state_machines.htm

[11] ChatMapper. Available: https://www.chatmapper.com/

[12] Articy. Available: https://www.articy.com/

[13] Key Archetypes of the Hero's Journey. Available: https://thenovelsmithy.com/8-archetypes-of-the-heros-journey/

[14] UC Berkeley. Monomyth: Hero's Journey Project. Available: https://orias.berkeley.edu/resources-teachers/monomyth-heros-journey-project