Answering Anya: A Virtual and Mixed Reality Game Centered on Nurturing

José Fuentes[1], Laura Restrepo[1], Boris Reyes[1], Ricardo Sánchez[1], and Pablo Figueroa[1][0000−0001−5412−8630]

Universidad de los Andes, Cra 1E #18A 40, Bogotá, CO
{jd.fuentes,l.restrepop,b.reyesr,r.sancheza,pfiguero}@uniandes.edu.co

Abstract. Virtual reality (VR) and mixed reality (MR) headsets offer a unique opportunity for immersive experiences, and coupled with the growing power of artificial intelligence, immersive realistic games can be developed. This project seeks to take advantage of these technologies and create a game centered around nurturing and the verbal interactions between a parent and her child, in MR platforms such as HoloLens 2, Meta Quest 3, Meta Quest Pro, and VIVE Pro Eye. We introduce Answering Anya, a game that aims to simulate realistic conversations with a virtual child, leveraging AI to generate questions and evaluate responses in a manner consistent with a young child's perspective. By integrating advanced features such as eye/gaze tracking and object detection, Answering Anya provides an engaging and challenging experience that enhances the player's empathy and communication skills. The game's ability to create meaningful and realistic interactions in both controlled and uncontrolled environments stresses its potential for fostering positive experiences and learning outcomes.

Keywords: Serious Games · Virtual Reality · Mixed Reality · Artificial Intelligence · Nurturing.

1 Introduction

The serious games industry is becoming progressively more influential, combining entertainment and education to address real-world issues. A topic that has not been explored enough in serious games is nurturing. Nurturing games focusing on parenting allow players to experience and learn from the challenges and rewards of raising a child. These games help improve parent-child interactions by fostering empathy and understanding of a child's needs, enhancing communication, and strengthening bonds through immersive experiences that extend beyond the screen. Hence, taking advantage of the immersive experience that VR headsets offer, makes for a fascinating chance to create a game where the player can truly feel like they are talking with a child.

We propose to explore novel technologies such as artificial intelligence and eye and gaze tracking in order to create a challenging and engaging serious game about nurturing.

2 Related Work

Currently, video games focused on this specific stage of nurturing are scarce. There are several games that touch on the topic, but they rarely center on the verbal interaction between a parent and their child. In addition, these games mainly involve babies who have yet to learn how to talk, or stages of swift growth where the children become teenagers or adults in a small window of time, leaving little to no room for the 3-5 age gap, such as in [2]. This scarceness is further highlighted when reducing the game platform to VR headsets, where the number of games is even smaller. This presents an opportunity to both create a game involving nurturing and adapt it to extended reality.

Games with interesting artificial intelligence capabilities have been presented in the past. For example Façade [1] involved users in a difficult conversation with a couple of friends. Games such as [3] also present players with difficult decisions that should be made when having kids. However, such conversations between the systems and the player are mostly textual and the games predate current developments in technology, which are now used in our game.

3 The Game: Answering Anya

The project's primary goal is to recreate a realistic interaction between a parent and her child regarding the questions the latter may ask. On the technical aspect, the main goal is to have a functional game working for the following devices: HoloLens 2, Meta Quest 3, Meta Quest Pro, and VIVE Pro Eye. These devices give a wide coverage of what it is the current state of the art in immersive technology. More specific goals include: using AI to generate imaginative questions and rate the answers of the player in such a way that it is coherent with a 5-year-old's worldview; integrating eye/gaze tracking and object detection components to determine the questions' subject; developing a fully functional game for a controlled (i.e., completely virtual) and an uncontrolled environment (i.e., superposed objects over the real world).

3.1 Environments

The game is presented in two main environments: the controlled environment refers to a completely virtual world that uses virtual reality, where the objects in the scene remain the same during each round, and where the player is fully immersed. On the other hand, the uncontrolled environment refers to the superposition of a handful of 3D objects over the real world using mixed reality, where the back cameras of the devices are used to show the player's surroundings, which can change depending on where the game is launched.

Both environments can be used in all devices we are targeting. However, the controlled environment is more suitable for virtual reality helmets while the uncontrolled environment is targeting augmented reality technologies and experiences. The following sections describe more details of both environments.

3.2 Game Flow

The game starts with a small introduction where the child's mom tells the player the instructions and explains that the girl's questions are causing her to abandon her drawing. The father answers her questions, which can lead to two outcomes: 1) a satisfactory answer that allows her to continue, or 2) an unsatisfactory answer that results in more questions. This cycle repeats until the girl finishes the drawing, time runs out, or a certain number of inadequate answers have been provided. The victory condition is met when the girl finishes her drawing within the available time.

3.3 Explored Technologies

Eye or gaze tracking is used in the controlled environment to track the player's eye or gaze position and trigger a question based on the object observed. For this, the particular devices' SDKs were employed.

Object detection technologies are used in the uncontrolled environment to label objects from the player's surroundings and trigger questions based on them. For this, the casting and tranmission capabilities of the devices were employed, paired with a local API that communicates with an external object detection service provided by Eden AI.

Artificial intelligence is used to generate the questions and rate the answers from the player, both from the point of view of a 5-year-old. For this, OpenAI's ChatGPT was employed.

Text to speech is used to generate an audio of the question formulated by the girl, which contributes to the overall game experience. For this, AWS' Amazon Polly was employed.

Speech to text is used to transcribe the player's answers into text, so they can be evaluated by the AI. For this, OpenAI's Whisper was employed.

Figure 1 shows how these technologies are assembled, in each one of the targeted immersive platforms. The system requires access to three main servers (i.e. OpenAI, AWS, and EdenAI through a local server with python code). The application combines results from these three servers plus an image capturing functionality in order to accomplish the desired gameplay. There are some limitations in some platforms; for example, the limited access of the passthrough data in the current Meta Quest Pro APIs. Nevertheless, at least one environment is currently functional in each immersive platform.

3.4 Interface Elements

Figure 2 shows the controlled environment. Notice the use of a TV as a board where the questions, answers and scores are shown, and an analog clock for the remaining time, which was designed this way to blend with the scene.

Figure 3 shows the uncontrolled environment. In this case, objects from the player's surroundings are shown. These objects can be used as part of the gameplay, thanks to the object detection capabilities. We add a white board and a

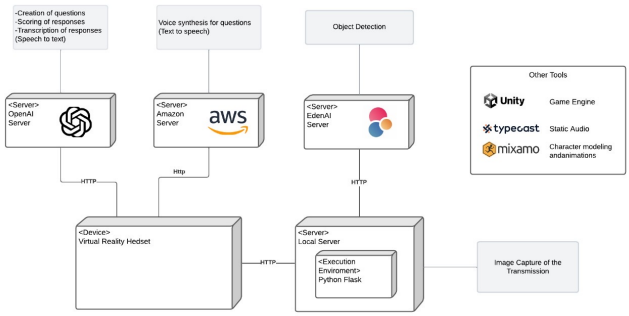

Fig. 1. Architecture of our solution

digital clock in order to show the state of the game, in the same way as in the controlled environment.

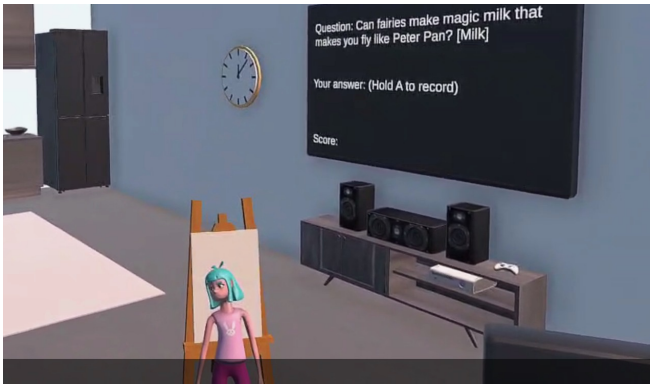

Fig. 2. Controlled Environment

Figure 4 shows the feedback provided by the child's drawing. Once a good answer is detected, the drawing advances to a new stage, which shows how the child is fulfilling her goal and at the same time the player is engaged with the questions that the child makes.

Additionally, Figure 5 shows some excerpts of the child's animation. Animations are meant to express feedback about the quality of a player's answer, in the form of the child's mood. Depending on the quality of the player's answer, the child could be happy, dissapointed, angry, or very sad, in that order.

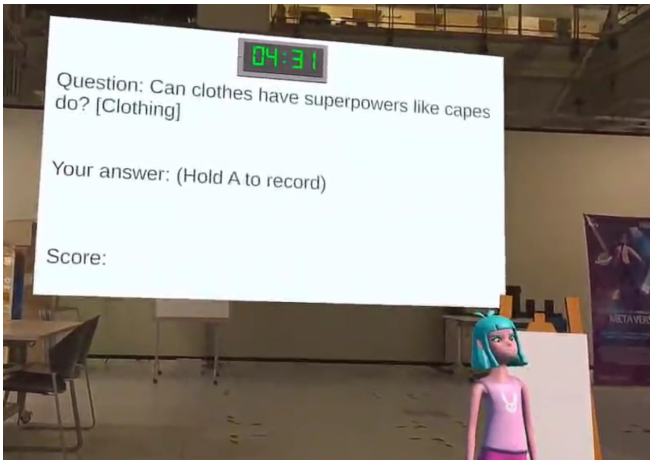

Fig. 3. Uncontrolled Environment

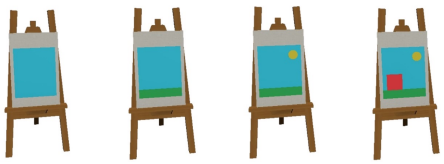

Fig. 4. Advance in the Child's Drawing

Fig. 5. Excerpts from the Child's Animation

4 Results

Two sessions of user testing took place to verify the game's usefulness and usability. The first one with the first working prototype was used to identify opportunities for improvement. From this test, we found out it was important to improve the initial description of instructions, add animations to the child according to the player's answers, show a textual version of the question so players could remember it better, and make changes to the child's model.

The second test was an open lab with the finished product, where new users were brought to experience the game and voice their opinions. We noticed that our changes solved the issues mentioned in the first test. Subjects also mentioned how interesting the concept of the game proved to be.

5 Conclusions and Future Work

We have presented Answering Anya, a multiplatform, immersive game that concentrates in the concept of nurturing. Several technologies are used in order to provide two versions of an immersive game, both for virtual reality and augmented reality users. The integration of artificial intelligence, eye tracking, object recognition, voice to text, and text to speech creates a compelling and powerful experience for users.

In the future, we plan to keep improving the prototype, given recent advances in these technologies, especially in artificial intelligence. A more throrough test will also help us to evaluate how the concept of nurturing is transmitted to users, and how it can be improved. Finally, novel devices such as the Apple Vision Pro could be explored, in order to evaluate the specific APIs and technologies available for that platform.

References

1. Playabl Studios, Façade. https://www.playablstudios.com/facade, last accessed 2024/07/07
2. Jason Rohrer, One Hour, One Life. https://onehouronelife.com/, last accessed 2024/07/07
3. Hosted Games, The Parenting Simulator. https://store.steampowered.com/app/1233720/The_Parenting_Simulator/, last accessed 2024/07/07