[EPICAST]

By

[Yueshen Li]

[Jianli Jin]

[Zihao Zhu]

[Di Wu]

Design Document for ECE 445, Senior Design, [Spring 2024]

Mentor: [Yushi Cheng] TA: [Guozheng He]

Abstract

"Dungeons & Dragons (D&D)" is celebrated for its boundless creativity and immersive gameplay yet faces limitations in visual and interactive expression of today's implementation. To address this, the "Epicast" is introduced, aiming to revolutionize the D&D gaming experience. Epicast integrates advanced technologies like GPT-powered AI and real-time data processing to offer players unprecedented immersion and engagement. The system features an AI Dungeon Master for intelligent narrative creation, an overhead projector for visual display, gesture recognition for intuitive control, and ambient lighting for mood setting. Epicast transforms traditional D&D sessions into dynamic, interactive adventures, enhancing player enjoyment and creativity. The design incorporates various modules such as Support, I/O, Control, Power, Game Instruction Decoder, and Game Logic State Machine, each fulfilling specific functions crucial to the system's operation. Challenges in hardware and software subsystems are addressed, ensuring compatibility, real-time data processing, and logical consistency. Ethical considerations and safety measures are also discussed to ensure a positive gaming experience for all users.

Contents

1. Introduction

Dungeons & Dragons (D&D) is a game that thrives on the breadth of imagination and the depth of interaction. It empowers players to construct elaborate worlds and characters that stem from the vast expanse of their creativity. This unfettered freedom not only offers a canvas for creation and interaction but also makes each gameplay experience profoundly personal and distinct. However, the richness of these imagined scenes and environments is often bottlenecked by the need for verbal expression and lacks an intuitive, sensory display. This limitation hampers the visualization of the game's full potential, presenting a hurdle for some players and constricting the game's broader allure. In addition, the Dungeon Master (DM) which is one of the players serving as the game's narrative architect, is often burdened with extensive preparatory work, juggling game mechanics with storytelling, which can be arduous and time-consuming. To address these issues above, project "Epicast" is a comprehensive system designed to enhance the Dungeons & Dragons gaming experience.

1.1 Functions

Central to our "Epicast" system is a GPT-powered AI that serves as an automated Dungeon Master, guiding gameplay with intelligent narrative creation and player interaction.

The visual aspect is handled by an overhead projector, capable of displaying intricate game scenes, animations, and simulating actions such as dice rolls directly onto the gaming surface. Its height is adjustable for optimal image quality. Gesture recognition is enabled through a sophisticated camera, allowing for intuitive control and the ability to capture memorable moments. Audio immersion is provided by integrated speakers and microphones for voice commands, narrative flow, and dynamic sound effects. Completing the sensory experience are ambient lights that adjust to the game's mood, providing synchronized lighting effects.

1.2 Benefits

Epicast transforms traditional Dungeons & Dragons (D&D) sessions into immersive, interactive experiences. Through the integration of advanced technologies like GPT-powered AI, sophisticated projection, and real-time data processing, Epicast opens new realms of imagination and engagement for players. Here's a scenario depicting the enhanced experience Epicast offers:

Imagine a group of friends gathered around a table, ready to embark on a D&D adventure. As they settle in, the lights dim and the Epicast system comes to life. The room is filled with the ambient sounds of a bustling medieval town, setting the stage for their quest.

With a wave of a hand, one player casts a spell. The gesture is instantly recognized by Epicast's sophisticated camera system, and the overhead projector vividly displays the spell's effects across the gaming surface – a brilliant burst of light enveloping their foes. Laughter and cheers fill the room, not just for the spell's success, but for the seamless, intuitive interaction that made it feel real.

The GPT-powered AI Dungeon Master, with its vast narrative capabilities, guides the adventure with dynamic storytelling. It generates rich descriptions of their surroundings, intricate plot twists, and detailed NPC interactions based on the players' actions and decisions. The players are no longer just imagining the dark, eerie forest; they see it projected in front of them, hear the sounds of unseen creatures, and feel the tension of the unknown.

As the adventure progresses, the ambient lights shift to match the mood, further drawing the players into the game. A serene blue glow accompanies their arrival at a tranquil lake, while a sudden shift to red signals the imminent danger of a dragon's attack. These cues, synchronized with the game's events, enhance the emotional and sensory engagement of the players.

With Epicast, every action and decision has an immediate, visible impact. The system's real-time processing and projection capabilities allow for a dynamic game world that evolves with the players' journey. The traditional boundaries of the board game are expanded, offering a more enriching and captivating experience.

1.3 Related Background

For a more immersive board game experience, several teams and companies have developed AR (Augmented Reality) board games for mobile devices such as phones or tablets. One notable example is the renowned game team known as "Mirrorscape,"[1] dedicated to crafting augmented reality platforms for tabletop role-playing games (TTRPGs). They have made available an AR D&D board game on both the Google and Apple app stores, accessible for free. Each player needs to position their device on a support with the cameras directed towards the table. Through their devices, they can witness virtual 3D game interfaces overlaid onto the real table. However, this implementation has its limitations. AR technology, as it stands, isn't flawless; there are instances where devices struggle to accurately recognize the table's plane, resulting in an unstable AR 3D interface. Additionally, players are required to view the interface through their device screens, which may detract from the gaming experience. Furthermore, with devices and supports occupying table space, players may find it uncomfortable to rest their arms.

1.4 Features

Our innovation, EPICAST, endeavors to enhance the gaming experience through various means. EPICAST sets up by requiring only one device for a group of players, thereby freeing up table surface space. Through camera detection, players can engage with the game using gestures, adding an interactive dimension. Moreover, EPICAST uses the power of AI LLM models to generate dynamic and enthralling gaming narratives and configurations--- the essential elements of D&D games. Our EPICAST system is committed to maximizing players' enjoyment. We firmly believe that EPICAST, as a comprehensive new board game system, offers distinctive and competitive features compared to existing products in the market.

1.5 Visual Aid

## 1.6 High-level requirements list

1. The EPICAST system functions as the "dungeon master", which is the judger of D&D game, should swiftly managing player inputs like voices and gestures, and also provide updated game status through projectors and verbal game status hints via speakers. Hence, EPICAST necessitates minimal delay in voice recognition, gesture detection, and game status processing.

2. EPICAST must accurately process the gathered data. Leveraging AI detection models, it translates players' gestures and speech into corresponding textual descriptions, subsequently transforming these texts into appropriate D&D game commands. EPICAST demands high-quality and precise interpretation of players' intentions.

3. Serving as the "dungeon master," EPICAST possesses the ability to shape game plots and make critical status judgments. Enhanced by LLM AI generation models, it ensures the coherence of game state logic, avoiding contradictions or ambiguities throughout the gameplay.

2 Design

2.1 [Block Diagram]
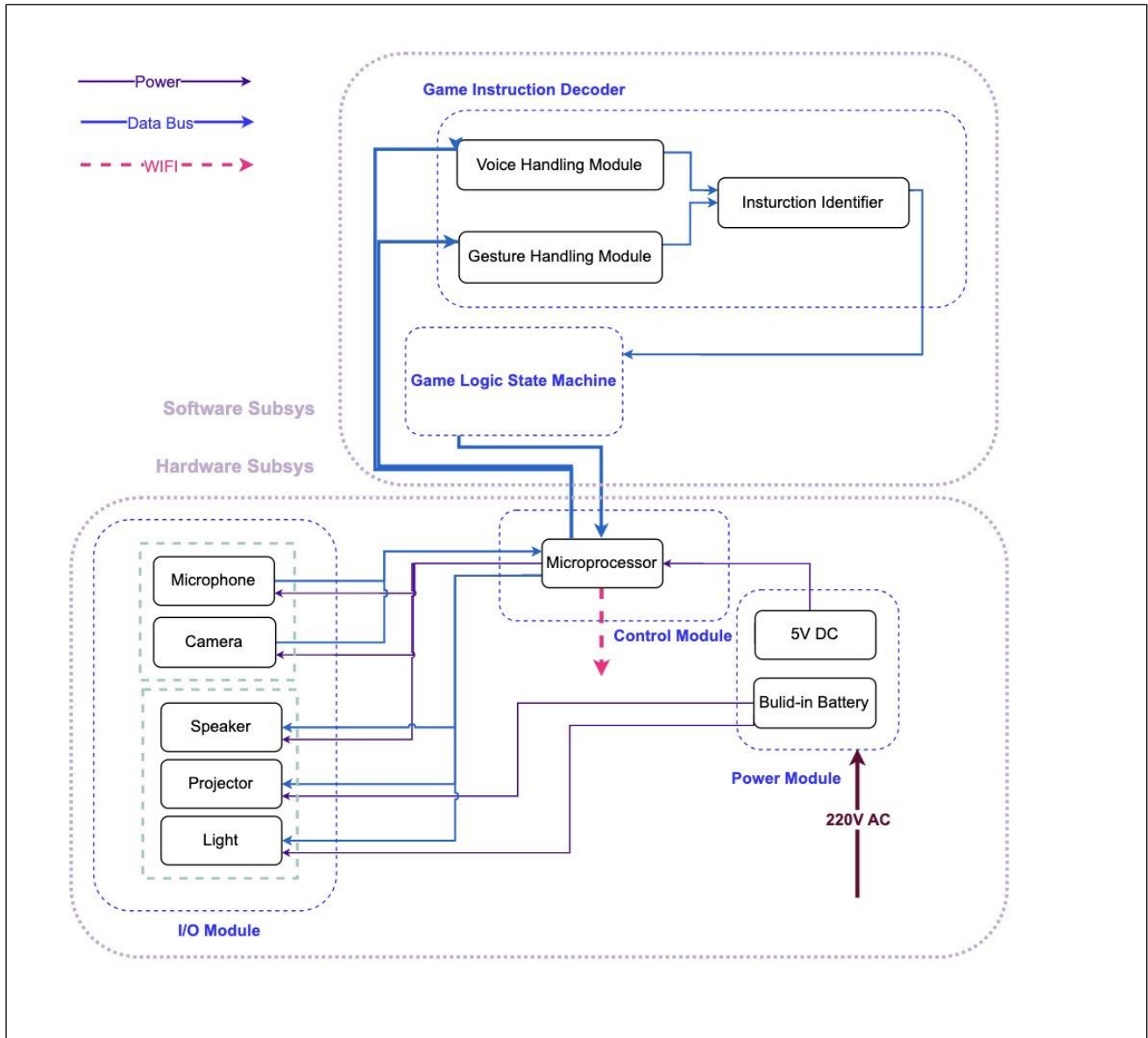
2.2 [Physical Diagram]
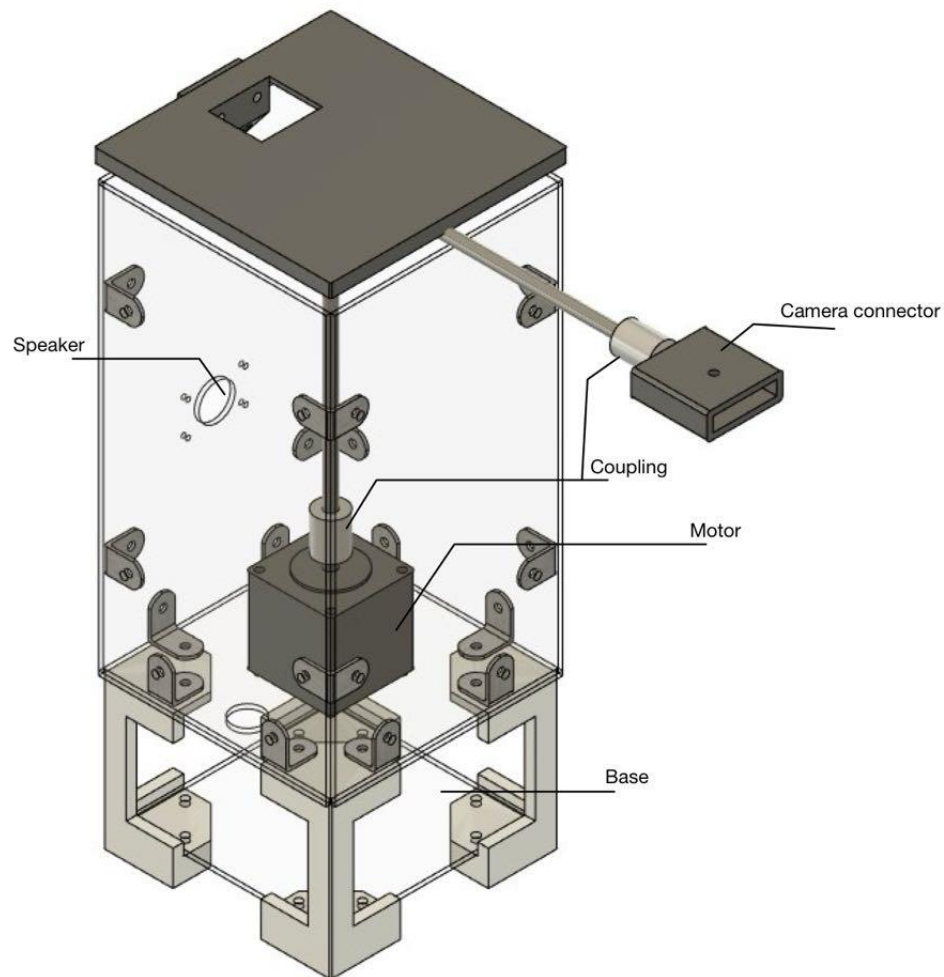
[Support Module]

2.3 [Block Description]

2.3.1 [Support Module]

The support module is a holder of physical parts like camera, microphone, raspberry Pi. During the game time, this device will be placed at the middle of the table, while four players sit at any position beside the table. When a player wants to manipulate his role in game, he just needs to reach out one of his hands into the sight of the camera. Once the camera detects player's hand, the rotating camera will stop, and the player can use gesture or voice to manipulate the role in game.

There will be a motor inside the device which can rotate the camera so that our camera can capture user's hands accurately. A camera with microphone will be attached to the support module by the camera connector shown in

the figure, The camera is lifted and slightly facing downward so that players can put their hands on the table instead of raising them towards the camera which can add more ease to their manipulation. A light will be placed near the lens so that it can point out the zone where the camera is capturing. A joint is used to connect the camera with the motor. With the joint we can slightly change the angle of the camera so that our device can fit the tables with different sizes.

Requirement:

1. Support Module can hold all physical parts including I/O Module, Motor Drive Module and Power Module tightly without bulk or break.

2. Support Module should dissipate heat well so that the inner electric components won't get too hot.

3. The light should point the destinated area accurately.

4. The motor can provide enough torque during the operation.

Verification:

1. Simulate parts to find out the quantity and location of the stress they suffered and the strain they occurred. Calculate the safety factor of the mechanism.

   After analysis, we believe the part which is likely to fail is the camera connector, since the protrusion of the stick will cause stress concentrations. Also, when adjusting the angle of the camera, an external bending will be applied on it. Given it's 3D printed part which are more likely to fail, it's necessary to make some simulations.
   In order to simply the simulation, I assume the connector will suffer a force of 30N downward which distributed normally and greater than the actual case.
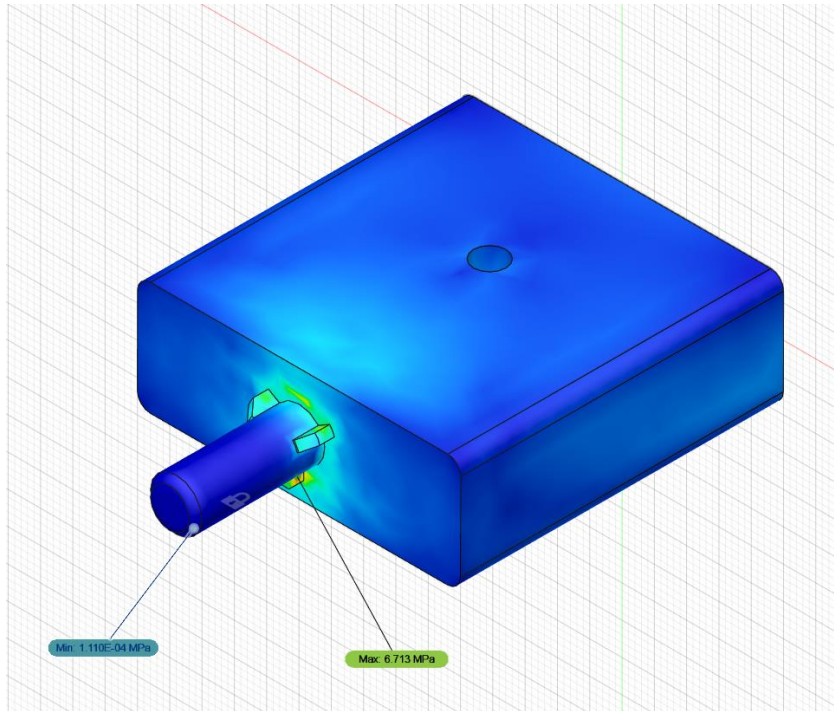
Figure. Von Mises stress suffered by camera connector.

As we can see from this figure, the stress is mainly contributed at the joint of the bar and the holder. In order to prevent this from failing, some reinforcement structures are added.
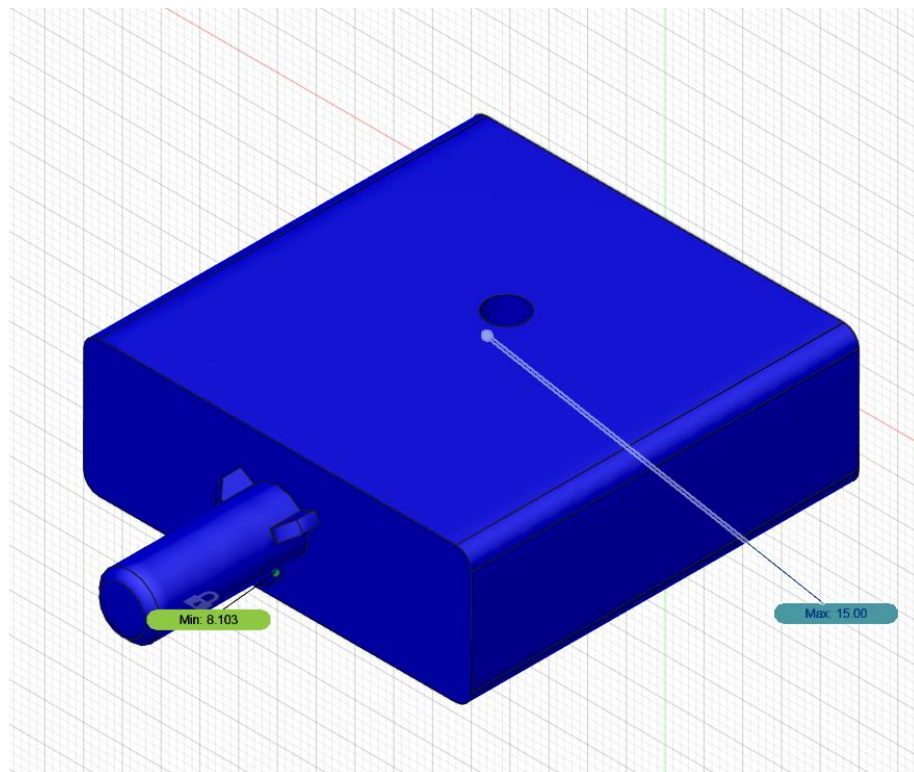

Figure. The safety factor of camera connector.

This holder has a minimum safety factor of 8.103, which satisfies the requirement of engineering.

2. Adopt hollow-out design for Support Module for better heat dissipation.
3. Compare multiple times through experiments, keep adjusting the position of the light until the allowable error range is reached.
4. Select the model of the motor reasonably and test it.

The model we select can provide a torque of 1.2NM.

Assume, the camera can be treated as a point mass.

Therefore：

The moment of inertia of camera: $I_1 = ml^2 = 0.3kg * 0.04m^2 = 0.012kg * m^2$ *(0.3kg is heavier than the actual case)*

The moment of inertia of horizontal bar: $I_2 = \frac{1}{3}ml^2 = \frac{1}{3} * 0.316kg * 0.04m^2 = 0.004kg * m^2$

The total moment moment of inertia r: $I = I_1 + I_2 = 0.016kg * m^2$

Assume, the acceleration of motor is a uniform acceleration motion. Therefore:
$$\tau = I * \alpha = I * \frac{\omega}{t} = 0.016 * 35 * \frac{2\pi}{60 * t} = 0.59N * m$$
(I assume the time needed to fully complete acceleration is 0.1s)
`

It turns out the motor we choose can fully satisfy the requirement theoretically.

2.3.2 [I/O Module]
The I/O module is a signal management and dispatch center. It connects with support module and software subsystem. The center of the Electrical subsystem is a Raspberry Pi broad. It receives real-time video messages from camera and microphone by Universal Asynchronous Receiver/Transmitter. After receiving the vision and audio signal, it will send them through Controller Area Network to software subsystem to accurately process. The application of virtual network cards will assist in online processing and operation of information and output. When specific images are output by software, it is also responsible for projector output signal transmission. Speaker is also applied to broadcast voice signal. For EPICAST, we have chosen the Raspberry Pi 4B/8G as our microprocessor. Serving as the software foundation, Raspberry Pi will operate on Linux, supporting the game instruction decoder and game logic state machine. The control module will also feature a data bus connecting all I/O components to facilitate interaction with players.

Choose to use a more integrated camera module instead of camera components. This makes the module output more stable and provides better mechanical strength with the help of packaging materials, reducing the need for strength considerations in later physical construction. The camera and microphone input are integrated in the component, which simplifies the component composition and makes the design smoother and more efficient.

Figure. Integrated camera and microphone.

To set up Raspberry Pi, begin by gathering the necessary equipment such as the Pi board, microSD card with an installed OS, peripherals like a keyboard and mouse, and a power supply. Install the OS onto the microSD card using software like Etcher, then insert the card into the Pi and connect peripherals. Boot up the Pi and follow on-screen instructions to set up the OS, including language, time zone, and network settings. Explore the Pi's capabilities for various tasks like programming or media center use. Learn the operating environment of the Raspberry PI. Through the configuration of the mirror system to achieve the basis of internal program editing, the Raspberry PI system learning and configuration process, the successful realization of the camera call, for gesture recognition and other visual input system to provide the foundation.

We chose a projector that fits needs, set it up, and connected it to my laptop and smartphone to test its functionality, including Bluetooth capabilities. After adjusting basic settings like brightness and focus, We assessed its performance, ensuring the projected image was clear and visible from different angles. Once satisfied, We finalized the setup, ready to use the projector for various purposes. The speaker also apply I2C signal from raspberry pi to recognize test from game machine and output host voice in short time.
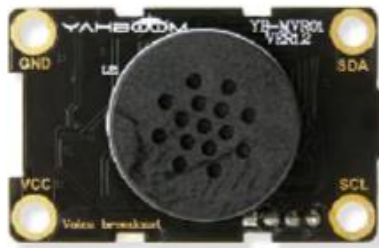

Figure. Projector.

Figure. Speaker.

2.3.3 [Motor Drive Module]

EPICAST image input initializes all player gesture information and uses the camera to capture round table player gestures in real time during gameplay for in-game interaction. The camera module will be equipped with a circular motion camera track for continuous monitoring. After a specific range of player gestures is obtained, the movement stops, and the player interaction is focused until the player completes the current turn. Supporting lighting system will improve the player experience in this process. To this end, a stable motor is needed to drive the camera mobile module to serve the player image input and recognition.

We choose N20 gear motor and TB6612FNG drive module with Raspberry PI to drive and control the motor. The motor drive and Raspberry PI 4B share a 5V 3A power supply.
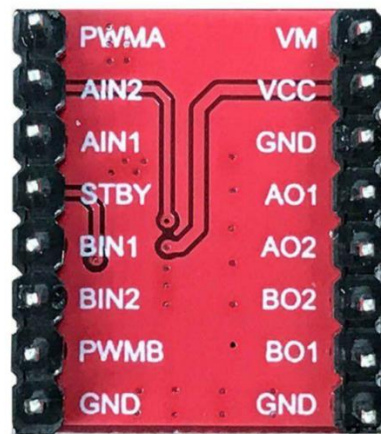
[Raspberry PI 4B and TB6612FNG Drive Module]
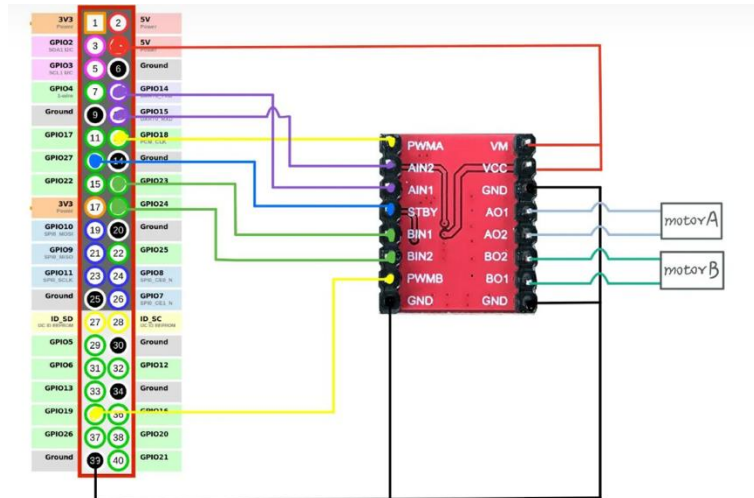
VM	Motor drive power input (2.74-10.8V)
VCC	Logic level input (5.0V)
AO1 and AO2	Connect to DC motor A
PWMA	Control the speed of motor A
AIN1 and AIN2	Control motor A (stop, forward, reverse)

STBY	Normal working/standby status control terminal
GND	Ground

*When AIN2 is true, the motor will forward work. WhenAIN2 is false, the motor will stop.

[Interface description of TB TB6612FNG]

Raspberry PI GPIO Port (BCM code)	TB6612FNG Driver Port	Description
27	STBY	Set the status of the TB6612FNG driver module (Working/Standby)
18	PWMA	Control the speed of motor A based on PWM
14	AIN1	Control motor A to stop, forward and reverse
15	AIN2	Control motor A to stop, forward and reverse

[Connections Between the Raspberry PI GPIO and the TB6612FNG]

Power subsystem should convert 220V AC input to 5V DC output. 5V DC is used to drive Raspberry Pi broad and TB6612FNG driver. Separate battery is supplied for projector. This part function properly with 220V AC power supply in China and concurrently supply 5V 3A with thermal stability.

2.3.4 [Game Instruction Decoder]
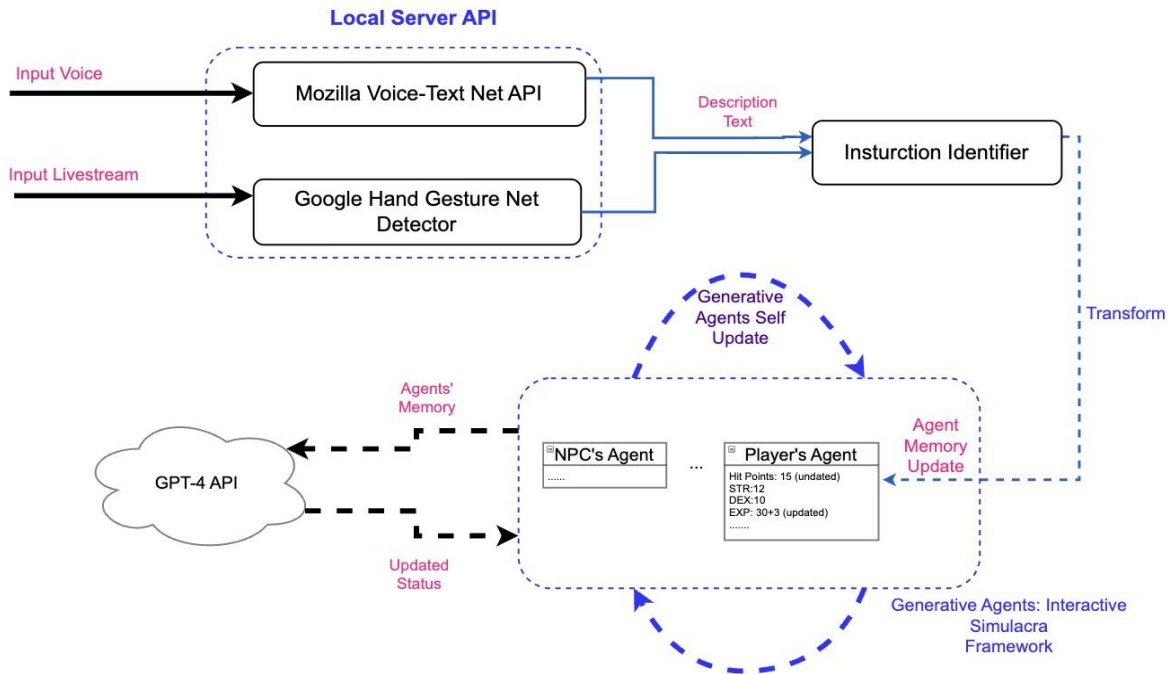
[Integrated Diagram of Decoder with GLSM]

The Game Instruction Decoder translates high-level voice and gesture commands from players into a set of clear instructions that the Game Logic State Machine (GLSM) can understand. It utilizes separate Deep Learning networks (Mozilla Speech-To-Text and Google MediaPipe Gesture Recognizer) to process sound and image data, respectively, and convert them into text. Both Mozilla and Google DL API are open-source tools and can be accessed through Internet. To improve the accuracy of "voice input to game instruction conversion," a keyword recognition system is also employed. The decoder should be able to:

1. Successfully recognize players' gestures from various angles

2. Process multiple players' gesture command simultaneously

3. Interpret player voice commands correctly despite variations like Inverted sentence order, Pauses and Rich player vocabulary.

4. Translate speed of players' voices should be within 2 words per second. The detection of one gesture should be within 1 second.

2.3.5 [Game Logic State Machine (GLSM)]

The GLSM acts as the system's core, managing the internal state of each player (items, health, stats) and the environment (map, NPCs, triggers). It ensures the game's logic flow remains consistent throughout the experience. The GLSM should be able to:

1. Guide and maintain the correct game progression despite variations in AI-generated content (e.g., narratives, scenarios).
2. Maintain low latency when contacting the server for the Large Language Model (LLM), ensuring smooth gameplay.
3. Modular design that allows for easy modification and adjustment to expand or change game content.

With careful analysis, we select the "Generative Agents: Interactive Simulacra of Human Behavior"

Upon thorough evaluation, we have adopted the "Generative Agents: Interactive Simulacra of Human Behavior" AI framework as the foundational structure for our Game Logic State Machine (GLSM). This model is selected for its capacity to maintain a long-term memory state for each character and to support intricate interactions among characters, thereby enhancing the dynamism and continuity of the game environment. This integration is visualized in the conceptual graph from the original paper (Park et al., 2023, Figure 1)."

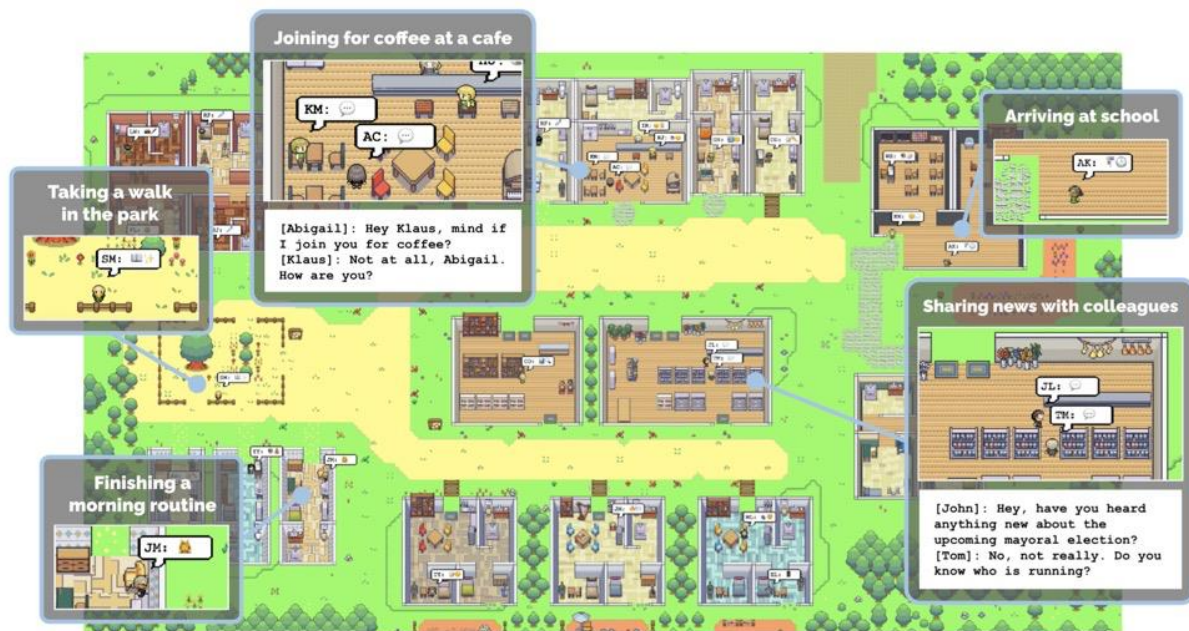Figure 1: Generative agents are believable simulacra of human behavior for interactive applications. In this work, we demonstrate generative agents by populating a sandbox environment, reminiscent of The Sims, with twenty-five agents. Users can observe and intervene as agents plan their days, share news, form relationships, and coordinate group activities.

2.4 Risks & Challenges

Physical Subsystem:

1. Lightweight and movable. The physical subsystem should be light and easy to handle with, which can make it fit more scenarios and can meet the needs of more customers, like can be played even in outdoors. This requires ingenious mechanism.
2. The other challenge is that our motor must be controlled precisely. Since our camera aims to put users' hand in the middle of its sight, but the camera is rotating, it can be hard for the motor to drive the camera to an appropriate angle. Also, there is a connector at one edge of the support, so we need to put a limit and find a path for the motor to follow to avoid it to hit the connector.

Hardware Subsystem:

1. Diverse sensor types. Different sensors may have different data formats, communication protocols and interface requirements. Hardware systems need to be able to adapt and handle a variety of different types of sensors, including microphone signal input and camera signal input. This requires hardware systems to be flexible and compatible, capable of effective communication and data exchange with a variety of sensors.

2. Data conflict and redundancy. When multiple information sources provide data at the same time, data conflict or redundancy may occur. Hardware systems need to be able to handle these conflicts and redundancies, selecting the right data for processing and analysis to provide accurate results. In addition, there are high requirements for real-time data. The hardware system needs to be able to process and respond to the input instructions of the sensor in time under the real-time requirement to ensure the stability and reliability of the system.

3. Stable motor drive. Because the moving camera module is an important part of the game interaction, the running speed and stability of the module will greatly affect the player interaction experience. The battery drive needs to be well matched with the identification output, improve the fluidity of movement, and complete the preset operation logic requirements to achieve practical goals.

Software Subsystem:

1. Maintaining Logical Consistency. A core challenge lies in maintaining the internal logical consistency of the game world despite its reliance on AI-generated content. The system must seamlessly integrate the AI-powered narrator's descriptions, visual elements generated by the projector, and game logic updates managed by the GLSM. This requires implementing deterministic code alongside the AI components. Deterministic code refers to pre-defined rules and algorithms that guarantee consistent and predictable outcomes, even when handling the inherent variability of AI-generated content.

2. Handling High Player Freedom: Dungeons & Dragons is renowned for its emphasis on player freedom, allowing players to make diverse choices and engage in creative actions. However, this presents a challenge for the system, as it needs to translate this vast range of player commands into a finite set of understandable game instructions. To address this, the Game Instruction Decoder and Game Logic State Machine together plays a crucial role.

3 Cost
3.1 Cost Analysis

Labor cost: The labor cost is an important part for the senior design and the cost are estimated as below.

Material cost: the material cost contains all subsystems listed below.

Part	Cost(¥)
Camera	259
2 Mic Array	89
Raspberry Pi 4B	889
Projector	299
Speaker	42
RGB Light	11.9

4 Schedule

Time	Yueshen&Jianli (Software System)	Di&Zihao (Mechanical Control)
Week 6	Finalize on gesture-control/text-to-speech/AI-generate-Image network	CAD the Mechanical components. Buy projector and camera, as well as raspberry.
Week 7	Try to reproduce the results of "Stanford town" Generative Agents: Interactive Simulacra of Human Behavior	Finalize the size of our mechanical components and the CAD.
By April 15th	Implement basic general game logic and transform "Stanford town" into combat scenario game. Include basic features as: health, attack value, map, turn-based etc..	Print and buy all of the needed parts. Assemble all the parts together Manage a suitable speed of the motor. Manage for an accurate capturing area of camera to place our flashlight.
Mid-test	Test for: 1. Character movement based on gesture control 2. Basic logic implementation for combat game Complete these tests on personal PC	Test for: 1.The tightness of our mechanical part can hold every component. 2.The smoothness and experience when manipulating the device,

By end of April	Further improve the game logic and add features, try to integrate GPT into game background generation	Fix any encountered bulks or unreasonable design. Try to improve the experience of users.
By the mid-May	Combine all parts of design and transplant software system to a Raspberry PI	Assemble all the parts together.

5 Ethics and Safety

Mechanic Part: We conduct a comprehensive risk assessment when designing mechanical equipment and prioritize the use of intrinsic safety measures to fundamentally eliminate the existence of hazards. We will add enough protective measurements to make sure the safety of our user like we have adopted a lightweight design to reduce the safety risk. These measures include enabling the machine to automatically prevent misoperation, as well as automatically detaching from the faulty area and safely transferring to the backup section or stopping operation in the event of a malfunction. We'll also fully consider the issue of human-machine matching to make the machine suitable for various human operations. We take the convenience of product installation into fully consideration.

Privacy Part: One paramount ethical consideration revolves around the privacy of user data within EPICAST. Utilizing cameras to detect player gestures and behaviors, and microphones to capture player voices, may raise privacy issues. The IEEE Code of Ethics underscores the significance of safeguarding others' privacy[2]. In our EPICAST implementation, while players' operation data is essential for game progression, we are committed to preserving privacy. Upon the completion of each game session, EPICAST will promptly delete the current game status and eradicate all residual players' data obtained from real-time voice and gesture inputs. Additionally, EPICAST transmits data to Google and Mozilla AI service platforms over the Internet, subjecting it to their respective privacy terms. Google and Mozilla have meticulously defined privacy protocols[3][4], thereby becoming our privacy policy dependence accordingly. In light of potential data theft during Internet transmission, we pledge to establish secure and encrypted channels for EPICAST's Internet access, ensuring the utmost protection of players' data privacy.

Artistic Copyright: With the implementation of AI in Epicast, we are entering a realm where creative expression and copyright laws intersect in new and complex ways. The generation of images and narrative content through AI systems raises important ethical considerations regarding artistic copyright. The use of such technologies must respect the original creators' rights and avoid infringing upon existing copyrighted works. Therefore, we have proposed the Ethics Code for AI-Generated Artistic Content:

(1)Respect for Originality: Ensure that all AI-generated content is sufficiently original and not a direct copy of existing copyrighted works. When inspiration is drawn from copyrighted materials, the AI should transform the inspiration in a significant way that results in a new, original piece.

(2)Transparency: Clearly disclose to all players that the visual and narrative content is generated by AI, and where feasible, provide information on the data sources used to train the AI models.

(3)Fair Use Compliance: Operate within the bounds of 'fair use' policy, where AI-generated content is used for commentary, criticism, reporting, or teaching in a way that does not diminish the market value of the original works.

(4)Continuous Monitoring: Regularly review and update the AI's programming to ensure ongoing compliance with copyright laws as they evolve, and promptly address any unintentional infringements.

(5)Creator Collaboration: Whenever possible, collaborate with artists and creators to obtain licenses for the use of their work as inspiration for AI-generated content, thus supporting the artistic community.

(6)Responsiveness to Concerns: Establish a protocol for responding to copyright holders' concerns about AI-generated content, including a clear process for reviewing and addressing potential infringements.

(7)Promotion of Ethical Use: Advocate for the ethical use of AI in artistic creations across the industry, setting a standard for responsible innovation that honors and upholds the rights of creators.

By adhering to these principles, Epicast will not only foster a fair and legally compliant gaming environment but also contribute to the responsible development of AI technologies in respect to artistic expression. The project will continuously strive to balance the innovative use of AI with the utmost respect for intellectual property rights, ensuring a positive and sustainable integration of technology and creativity.

References

[1]   Mirrorscape Team. "Introduction to Mirrorscape Team." Mirrorscape, www.mirrorscape.com. Accessed: March 5, 2024.

[2]   IEEE, "IEEE Code of Ethics," June 2020. [Online]. https://www.ieee.org/about/corporate/governance/p7-8.html. Accessed: March 7, 2024.

[3]   Google. "Google Privacy Terms," [Online]. Available: https://policies.google.com/privacy?hl=en-US. Accessed: March 7, 2024.

[4]   Mozilla. "Mozilla Privacy Terms," [Online]. Available: https://www.mozilla.org/en-US/privacy/. Accessed: March 7, 2024.

[5]   Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative Agents: Interactive Simulacra of Human Behavior. arXiv. https://doi.org/10.48550/arXiv.2304.03442