Breathing Life into Animatronics

JORGE DAVO SAINZ

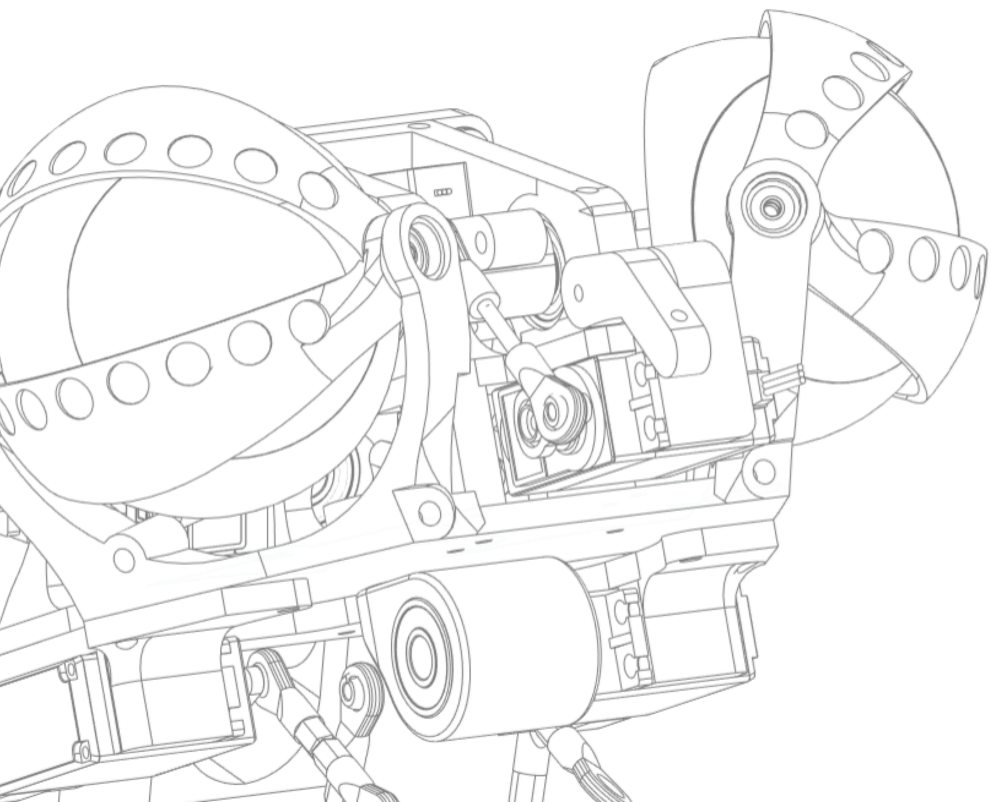
CREATIVE TECHNOLOGY

BACHELOR THESIS

UNIVERSITY OF TWENTE

SUPERVISOR: **EDWIN DERTIEN**

CRITICAL OBSERVER: **ROBBY VAN DELDEN**



This thesis is dedicated to my incredible family, whose love and support have fuelled my passion for creating, even while since a short age my projects might have seemed a little crazy... and they still do. Thank you, Isa, Mamá y Papá for always believing in me and encouraging me to make my own path.

To the phenomenal team at P&P Projects, thank you for sharing your expertise, your creativity, and your commitment to pushing the limits of themed entertainment. Your collaboration and friendship have been invaluable in bringing this project to life.

I extend my deepest gratitude to my supervisors at the University of Twente, Edwin Dertien and Robby van Delden, for their guidance, insightful critiques, and continuous support throughout this crazy project. Your mentorship has been instrumental in shaping my research and helping me grow as a creative technologist.

Table of Contents

1 Introduction

The theme park landscape is evolving, moving towards experiences that are dynamic, interactive, and personalized [1]. Animatronics, captivating figures that blur the lines between fantasy and reality, are a key aspect of this transformation. While animatronics have captivated audiences with their lifelike movements and intricate details, the industry is now looking to elevate their capabilities beyond pre-recorded sequences [2]. The demand is for animatronics that can engage in natural interactions, responding to their surroundings and the people they encounter in a way that feels spontaneous and genuine.

This shift towards interactive animatronics reflects a broader trend in society, where digital technology has fostered a growing expectation for immersive and engaging experiences[3]. Video games, mobile apps, and virtual reality platforms have all contributed to this phenomenon, showcasing the potential for personalized narratives and characters that react dynamically to user input [3]. P&P projects recognizes the need to adapt to these evolving preferences, seeking to offer visitors experiences that are not only visually pleasant but also emotionally resonant and intellectually stimulating.

Animatronics offer a unique platform for achieving this goal. Unlike their digital counterparts, they exist in the physical world, allowing for a tangible connection with audiences[4]. The challenge lies in bridging the gap between the mechanical precision of animatronic movements and the fluid, natural interactions of living beings.

This is where the concept of a multi-layered interaction engine comes into play. To create the illusion of responsiveness, we need a system that can create and blend between a variety of behaviors, adding subtle background movements with more deliberate actions and reactive responses. This layered approach is key to achieving the natural interaction this project aims to achieve.

This project is a collaborative effort between P&P Projects and me, P&P projects is a leading company in the creation of themed elements for attractions and experiences. P&P Projects has consistently pushed the boundaries of what's possible in the realm of themed entertainment, and their commitment to innovation aligns perfectly with the goals of this project.

The antecedents of this project lead back to 2023 with the development of a highly advanced animatronic creature, containing 21 degrees of freedom for incredibly detailed and lifelike movements. This creature serves as the perfect platform for exploring the potential of a multi-layered interaction engine, and its creation has laid the groundwork for the next phase of the project: bringing the animatronic to life through sophisticated interactions.

Figure 1: Animatronic Creature.

I was part of this project as an intern where my role focused on the mechanical design, electronics, programming, and assembly. P&P Projects oversaw the design and construction of the creature [Figure 1], contributing their expertise in areas such as character design, painting, and theming.

The aim of this graduation project is to create an interaction engine that when combined with the animatronic creature can create a natural interaction with theme park visitors, responding to their facial expressions, emotions, and even their proximity. Ultimately, this project aims at pushing the boundaries of what is possible with animatronic technology. By demonstrating the potential for natural interactions, it aims to inspire further innovation in the field and pave the way for a new generation of animatronic characters that captivate audiences and overall improve the theme park experience.

1.1 Research question

The project aims to create a multi layered interaction engine. Such an engine will serve as the animatronics' "brain," processing information from various sources and using it to orchestrate realistic responses.

The main research question: *"How can a multi-layered interaction engine be designed and implemented for an animatronic creature to enable natural interactions with theme park visitors?"* Will be the main research question that will guide the project, and it will be supported by the following sub-questions:

1. *Which current interaction architectures exist which accomplish an interactive behavior in other animatronic or robotic characters?*
2. *How have "vital signs" such as breathing and other subtle movements been implemented in other robotic characters to simulate lifelikeness?*
3. *What evaluation methods can be used to assess the interaction engine?*

1.2 Method

The main research question will be answered in three phases: theoretical exploration, realization, and a final product.

The theoretical exploration initial phase will focus on establishing an understanding of existing interaction architectures and techniques for simulating lifelike behaviors in animatronic and robotic characters.

To answer sub-question 1, a review of relevant literature will be conducted. Particular attention will be given to projects that have successfully implemented interactive animatronics or robots, analyzed the underlying architecture of their interaction engines, and identified key components and design principles.

Sub-question 2 will be addressed through a combination of literature review and the realization stage. Existing research on the role of subtle movements in conveying lifelikeness will be explored.

Sub question 3 will be answered in the evaluation phase, where the evaluation methods performed in the interaction engine will be described and tested.

Following the initial research phase, the project will transition into a realization stage where the theoretical knowledge will be applied to the development of the interaction engine. This phase will involve prototyping and testing. This stage will consist of building a functional prototype of the interaction engine. The prototype will integrate a newly designed interaction engine based on the investigation with the necessary software components. It will then be implemented into the existing creature animatronic.

The culmination of the research and realization phases will lead to the development of the final prototype of the multi-layered interaction engine. This prototype will represent the most refined and optimized version of the system, incorporating all the insights and improvements gleaned from the design process.

The final prototype will integrate the designed interaction engine with the animatronic creature's hardware components. This includes the communication between inputs, data processing units, and animation control systems. Finally, the engine's algorithms and parameters will be optimized to achieve a high level of responsiveness, realism, and interactivity.

This page has been intentionally left blank.

2 Background research

To lay the groundwork for the development of a multi-layered interaction engine, an exploration of existing research and technologies was conducted. This background research was guided by two key sub-questions that dive into the core challenges of creating interactive and lifelike animatronic characters:

Existing interaction architectures: which current interaction architectures exist which accomplish a similar behavior in other animatronic or robotic characters? This exploration's goal was to identify established models and design principles that could inform the development of a new interaction engine.

Simulating lifelikeness: how have "vital signs" such as breathing and other subtle movements been implemented in other animatronic characters to simulate lifelikeness? Understanding existing techniques for creating realistic and engaging behaviors was crucial for developing a new approach for an interactive engine.

2.1 Current interaction engines

Several approaches to designing and implementing interactive animation engines have been explored in previous research:

2.1.1 Layering laban effort features

One approach, proposed by Knight and Simmons [5] focuses on layering Laban Effort features onto pre-existing robot task motions. The Laban Effort System, originating from dance and acting training, provides a vocabulary for describing the dynamic qualities of movement, such as the use of space, time, weight, and flow. It becomes possible to fill a robot's movements with expressivity and convey different internal states or attitudes based on these features [5].

Knight and Simmons [5] demonstrated this approach by applying Laban Effort features to the head movements of a Nao robot during a "look-for-someone" behavior [Figure 2]. They found that certain features, such as the use of space and time, were recognizable by human observers and influenced their perception of the robot's state. This suggests that layering Laban Efforts can be an effective way to improve the expressiveness and communicative abilities of robotic characters [5].

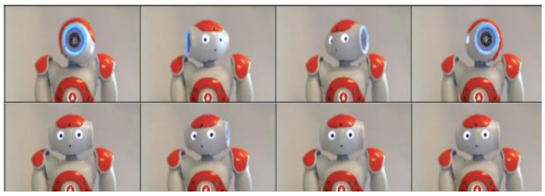


Figure 2: Nao during look for someone [25].

2.1.2The EMOTE model

Another approach is the EMOTE model, developed by Badler et al.[6]. EMOTE uses high-level parameters to control the animation of human characters, focusing on aspects such as posture, gesture, and facial expressions. The model allows for the specification of complex behaviors through a combination of pre-defined actions and real-time adjustments based on environmental factors or user input [6], [7].

2.1.3 Behavior markup language

The Behavior Markup Language (BML) offers a standardized way to describe and synchronize the behavior of interactive characters, including both verbal and non-verbal actions [8]. BML scripts can be interpreted by behavior realizer engines, such as AsapRealizer, which translate the high-level descriptions into platform-specific commands for controlling the character's movements, speech, and other outputs [9], [10]. AsapRealizer also incorporates feedback mechanisms to ensure smooth and synchronized execution of behaviors, considering the timing constraints and capabilities of the character's embodiment [9], [10].

2.1.4 Heterogeneous multilevel multimodal mixing

Davison et al. [11] introduced the concept of Heterogeneous Multilevel Multimodal Mixing (HMMM) as a comprehensive framework for orchestrating the complex interplay of behaviors in interactive robots. HMMM recognizes that truly engaging and lifelike interactions require more than just pre-programmed sequences or simple reactions to stimuli [11] Instead, it advocates for a harmonious blend of three distinct layers of behavior:

Deliberative behaviors: these actions are driven by a dialogue manager, which plans and sequences behaviors based on the current context of the interaction, the goals of the robot, and the perceived intentions of the user [11]. Examples of deliberative behaviors include initiating greetings, providing instructions, or expressing emotions in response to specific events [11].

Autonomous behaviors: these actions operate independently of the dialogue manager, often running continuously in the background to simulate lifelike qualities and maintain a sense of presence [11]. Examples of autonomous behaviors include breathing motions, blinking, and subtle shifts in posture or gaze direction [11].

Reactive behaviors: these actions are triggered in response to unexpected events or stimuli in the environment [11]. Examples of reactive behaviors include turning to look at a sudden movement, expressing surprise at a loud noise, or recoiling from an unexpected touch [11].

HMMM emphasizes the importance of seamless integration and prioritization of these different behavioral layers [11]. As seen in figure 3, the system must decide which behaviors to execute at any given moment, making sure they complement each other and create a natural flow of interaction [11]. This may involve suppressing autonomous behaviors when deliberative actions take priority or blending reactive responses with ongoing autonomous motions to maintain a sense of continuity [11].

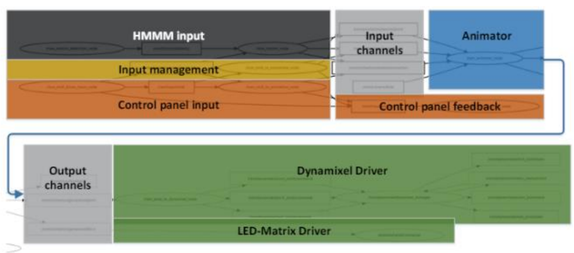


Figure 3: HMMM final system [11].

HMMM offers a promising approach for achieving lifelike interactions with robotic characters. It allows for the creation of robots that can exhibit a sense of autonomy and personality.

2.1.5 Parameterized facial animation

Wittig et al. [12] explored the use of parameterized facial animation for socially interactive robots [12]. Their approach uses Russell's circumplex model of emotion [13], which maps emotions onto a two-dimensional space based on arousal and pleasure [13]. By manipulating these parameters, the system can generate a wide range of facial expressions that reflect the robot's internal state or respond to external stimuli.

2.1.6 Key insights taken from the existing interaction engines.

Several key insights emerge from the exploration of previous research on interactive animation engines and will be implemented in the newly made interaction engine.

Importance of expressivity: whether through Laban Effort features, facial expressions, or other means, filling characters with expressivity is very important for creating engaging and believable interactions.

Multi-layered approach: effective interaction engines usually have a multi-layered approach, combining deliberative, autonomous, and reactive behaviors to achieve a balance between responsiveness and realism.

Feedback and synchronization: feedback mechanisms and synchronization techniques are very important for smooth transitions between behaviors which make it appear more natural.

Modularity: modular system architectures facilitate the development and adaptation of interaction engines which could be applied to different characters.

2.1.7 Summary of interaction engines

Based on the reviewed literature, the functionality of an interactive animation engine can be divided into three key stages: perception, decision-making, and action.

Perception: the engine gathers information about the surrounding environment and the state of the character itself. This data comes from various sensors, such as cameras for facial recognition and object tracking, microphones for speech recognition, or touch sensors for physical interaction. Internally, states like the character's current emotional state or energy level are also considered.

Decision-making: the engine analyzes the perceived information and determines the response. This is done by interpreting sensor data, evaluating the character's internal state, and applying decision-making algorithms to blend different base animations with external input. The decision-making process could be based on logic decision-making, artificial intelligence techniques, or a combination of both.

Action: the engine triggers the chosen action, for example: physical movements, facial expressions, speech, sounds or other kinds of output. The action should be executed smoothly and seamlessly, considering the character's current posture, the timing of previous actions, and the overall context of the interaction.

2.2 Gaze as a method for interaction

Gaze plays a significant role in establishing joint attention, which is the shared focus of two or more individuals on a common object or event [14]. Robots can use gaze following to identify the user's focus of attention and then direct their own gaze towards the same object, signaling their engagement and understanding of the interaction [14]. Moreover, robots can use referential gaze to guide the user's attention towards specific objects or locations, particularly useful in providing instructions or during collaborative tasks [14].

Gaze serves as a crucial cue in regulating turn-taking during conversations. Observing the gaze patterns of their counterparts allows humans to infer when it is appropriate to speak or listen. Similarly, robots can use gaze cues to show their intention to hold or give the floor during interactions [15]. For instance, avoiding gaze can indicate the robot is still processing information and intends to continue speaking, while directing gaze towards the user at the end of a sentence can signal the robot is ready for the user to respond [15].

One example of this, explored the use of animation principles to design robot gaze behavior, using elements like jerking and variations in the speed and timing of eye and head movements to improve the believability of the robot character [16].

2.3 Common input methods and sensors for robot interaction

Interactive robots rely on inputs through sensors interpret their environment and the people around them. These inputs allow the robots to respond dynamically to user behavior and engage in natural interactions.

2.3.1 Facial tracking

Facial tracking systems are the eyes of robots, allowing them to interpret non-verbal cues shown through facial expressions and gaze direction. These systems typically use cameras to capture video input, which is then analyzed by computer vision algorithms to identify and track facial features, such as the eyes, nose, and mouth [17].

Facial tracking data can be used for a variety of purposes. For example, Sanjaya et al. [17] developed a social robot that could recognize and imitate four different facial expressions (smile, sad, angry, and surprise) based on real-time facial tracking. Similarly, Pan et al. [16] developed a system for lifelike gaze behavior in an animatronic bust, using facial tracking to identify persons of interest and select gaze behaviors based on their actions and proximity. Facial tracking data can also be used to estimate emotional states, detect engagement levels, and personalize interactions based on individual user responses [17].

2.3.2 Emotion recognition

Emotion recognition systems identify and interpret the emotional state of the user, providing robots with a deeper understanding of human behavior and allowing them to respond with empathy and sensitivity [17]. Emotion recognition can be achieved through various methods, including analyzing facial expressions, vocal intonation, and physiological signals [17].

2.3.3 Sound

Sound input provides interactive robots information about their surroundings and the people within them. Microphones capture audio signals, which are then processed using sound recognition algorithms to identify specific sounds, such as speech, laughter, or crying, as well as to detect changes in the acoustic environment [18].

Sound recognition allows robots to respond interactively, engage in conversations, and help in various situations. For example, Kim et al [18] developed a real-time sound recognition system for a human care robot that could recognize custom sound events, like personalized doorbells and phone ring tones, allowing the robot to provide reminders and assistance to elderly users [18].

2.3.4 Touch

Cang [19] investigated the use of touch, gaze, and biometric data for emotion recognition in a therapeutic robot context. By analyzing touch gestures, gaze patterns, and physiological signals such as skin conductance and respiration rate, the system could distinguish between four different emotional states: stressed, relaxed, excited, and depressed. Emotion recognition systems have the potential to revolutionize human-robot interactions, enabling robots to provide emotional support, adapt their behavior to the user's mood, and personalize experiences based on individual emotional responses [19].

Cang et al. [19] explored the use of a custom-built fabric touch sensor for gesture recognition in a social robot context. The sensor, can detect both pressure and location, allowing the robot to distinguish between various touch gestures, such as stroking, patting, and rubbing, laying the groundwork for more responsive interactions based on touch input [19].

2.4 Animation blending

Animation blending is a critical technique for creating smooth and natural transitions between different animatronic states and behaviors. Without it, movements would appear jerky therefore breaking the illusion of life.

Radial Basis Function (RBF) interpolation offers a powerful tool for animation blending [20]. RBF interpolation utilizes a set of base functions placed in parameter space to approximate the desired blend weights for each animation. This technique can produce smooth and continuous transitions while also allowing for fine-tuning of the blending process to accommodate specific requirements or constraints [20].

Blend shape animation offers a popular approach for creating facial expressions, using a set of pre-defined shapes that represent different facial movements [21]. By blending these shapes together in varying proportions, animators can create a wide range of expressions. Joshi et al. [21] proposed a method for automatically segmenting blend shape models into meaningful regions, allowing for more intuitive and efficient control of facial expressions [21]. Similarly, Seol et al. [22] developed a technique for simplifying and editing blend shape weight graphs, enabling animators to refine and improve facial animations easier [22].

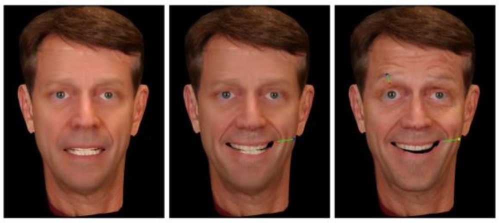


Figure 4: Animation blending in faces [21].

2.5 Current interactive robots on theme parks

While the concept of interactive animatronics holds immense potential, its practical implementation within theme parks remains relatively limited. However, a few notable examples show the possibilities for future development.

2.5.1 Disney's droids in Star Wars: Galaxy's Edge

"Disney's Star Wars: Galaxy's Edge has pushed the boundaries of immersive entertainment, transporting guests to a galaxy far away through a combination of cutting-edge technology and captivating storytelling" [23]. "Among the park's most innovative features are the interactive droids

Figure 5: Disney's Droids [23].

that roam the land, engaging with visitors and adding to the overall atmosphere of a bustling spaceport” [23].

These "BD units" or "buddy droids" are equipped with advanced robotics and sensors, allowing them to navigate the environment independently, respond to sounds in real-time, and interact with guests and other animatronics [23]. Their movements are AI based, providing a repertoire of actions such as dancing, head tilts, and even expressive eye movements [23]. The droids' ability to interact autonomously marks an important step forward in animatronic technology, showing the potential for creating characters that feel truly alive and responsive [23].

2.5.2 Disney's vyloos in guardians of the galaxy- mission: breakout!

Another example of interactive animatronics can be found at Disney California Adventure Park, where the “mischievous Vyloos from the Guardians of the Galaxy films reside within The Collector's Fortress” [24]. These small, alien creatures are housed in a pod and are powered by an onboard system that allows them to operate autonomously [24].

Figure 6 Disney's Vyloos[24] .

The Vyloos are equipped with sensors and cameras, allowing them to perceive and respond to guests' presence and non-verbal cues [24]. They have distinct personalities, from shy to outgoing, and their behaviors are dynamically adjusted for different interaction [24]. The Vyloos also show the potential for creating connections between visitors and animatronic characters.

2.6 Conclusion of background research

The exploration of existing research on interactive animation engines and related technologies has provided valuable insights and laid a solid foundation for addressing the central research question of this project: How can a complex, multi-layered interaction engine be designed and implemented for an animatronic creature to enable natural interactions with theme park visitors? The findings from the background research will contribute to answering this question and its associated sub-research questions.

2.6.1 First sub-question: existing interaction architectures

The review of current animation engines revealed several promising approaches that offer inspiration and potential solutions for designing a new multi-layered interaction engine. The concept of layering Laban Effort features onto existing motions highlights the importance of expressivity and the potential for conveying internal states and emotions through subtle movement variations [25]. The EMOTE model showcases the effectiveness of using high-level parameters to control character animation and provides a framework for building complex behaviors [26]. The Behavior Markup Language (BML) and behavior realizer engines like AsapRealizer offer standardized tools for specifying and synchronizing character behaviors, simplifying the integration of different interaction components [8], [9], [10].

The concept of Heterogeneous Multilevel Multimodal Mixing (HMMM) proposed by Davison et al. [11] is particularly relevant to this project’s goal. HMMM's emphasis on blending deliberative, autonomous, and reactive behaviors aligns perfectly with the vision for creating an animatronic creature that can respond dynamically to its environment and engage in natural interactions with visitors [11]. Additionally, HMMM's consideration of multi-party interactions provides guidance for developing an engine that can manage the complexities of real-world theme park scenarios.

2.6.2 Second sub-question: simulating lifelikeness

The investigation into how "vital signs" and subtle movements contribute to lifelikeness revealed the importance of going beyond simply mimicking human actions. The work of Ribeiro and Paiva [7] on applying animation principles to robots highlights the role of exaggeration, timing, and secondary actions in creating believable and engaging characters. Their research emphasizes the need to consider the limitations and unique capabilities of the animatronic platform when designing lifelike behaviors [7]. The exploration of existing animation engines also provided insights into techniques for implementing "vital signs." The use of background animations for breathing or subtle body movements [11], [14], [15], [16] demonstrates the effectiveness of these methods in improving the perception of life and presence.

2.6.3 Next steps

The background research has established a strong foundation for the development of a new multi-layered interaction engine. The next stage of the project will involve translating these theoretical insights into practical implementations in the realized interaction engine.

This page has been intentionally left blank

3. Methods and techniques

This chapter describes the approach for designing the multi-layered interaction engine into the animatronic creature.

3.1 Design phases

The development of the interaction engine was structured into three phases:

3.1.1 Conceptual design

The first stage focused on defining the architecture of the interaction engine, getting inspiration from the existing models explored in the background research. The main decisions were selecting the appropriate technologies for perception, decision-making, and action. After that the mechanisms for integrating and prioritizing various behavioral layers were also selected.

3.1.2 Prototype development

Based on the conceptual design, a prototype of the interaction engine was developed. This prototype included the necessary hardware and software components, including inputs for gathering environmental data, a logical unit for analyzing information and making decisions, and an animation control system for driving the creature's movements.

3.1.3Final implementation

The final phase involved integrating the refined interaction engine with the animatronic creature and testing for reliability and the set of requirements.

3.2 Technical approach

The technical implementation of the interaction engine involved a combination of hardware and software:

3.2.1 Perception system

Facial recognition: a camera-based system was used to capture and analyze visitors' facial expressions. This was done using existing facial recognition software: OpenCV.

Audio input: a microphone was used to capture audio input from the surrounding environment. This input was used for analyzing the overall level of noise and activity in the area and triggering responses from the creature.

3.1.2. Decision-making system

The decision-making serves as the brain of the interaction engine, analyzing perceived information and orchestrating the animatronic creature's responses based on a multi-layered architecture inspired by the principles of HMMM [11]. This architecture combines deliberative, autonomous, and reactive behaviors, so there is a balance between planned interactions, lifelike autonomy, and dynamic responses to unexpected events.

A rule-based system forms the foundation of the decision-making process, providing a framework for mapping perceived inputs to corresponding actions. These predefined rules define the creature's responses based on combinations of input data, internal states, and contextual factors.

Autonomous behaviors were layered onto the rule-based system to improve the creature's lifelikeness and create a sense of presence even when not directly interacting with visitors. These behaviors operate independently, running continuously in the background.

Reactive behaviors were integrated into the decision-making system to enable the creature to respond dynamically to unexpected events or stimuli in the environment. These behaviors will be triggered by specific inputs and changes in the environment.

The decision-making system considers the overall context of the interaction when prioritizing behaviors. For example, if the creature is in the middle of an interaction sequence with a visitor, it may be less likely to respond to minor environmental stimuli or changes in its peripheral vision.

3.1.3. Action system

The animatronic creature's movements are based on a keyframe looping animation mapped to different variables, the most viable option for this is Touch Designer. Other options such as Blender, Unity and Autodesk Maya have been considered, but Touch Designer was selected because of its flexibility and graphical approach.

Animation blending and queuing were implemented to make sure there are smooth transitions between different actions and maintain the illusion of lifelike movement.

3.3 Evaluation

Evaluating the interaction engine and incorporating expert opinions are vital parts of the development process, as they provide valuable insights for refining the system and making sure it meets the desired objectives. Feedback from both controlled testing and experienced professionals can uncover unforeseen problems, highlight areas for improvement, and validate design choices.

3.3.1 Prototyping and testing

The prototype of the interaction engine was tested in a controlled environment, simulating various interaction scenarios with the animatronic creature, and observing its responses. This allowed for the identification of any technical issues or unexpected behaviors, as well as the measurement of metrics that can evaluate the engine’s performance.

3.3.2 Expert evaluations

Feedback was also be gathered from experts in animatronics and in the theme park industry. Their insights provided perspectives on the technical aspects of the system, as well as its ability in achieving the desired level of realism and engagement.

3.4 Ethical considerations

The development of interactive animatronics raises important ethical considerations that must be addressed. Particular attention will be given to the following aspects:

3.4.1 Privacy

The system was designed to not store personal data from visitors. Facial recognition data was used solely for the purpose of driving the creature's responses in real-time, however there are still concerns regarding facial tracking technology and regulations.

3.4.2 Transparency

Clear information was provided to the users interacting with the creature about the nature of the animatronic creature's interactive capabilities. This information explained how the system works and what data is collected.

4. Stakeholder identification and analysis

Developing a multi-layered interaction engine requires an understanding of the needs of all stakeholders involved.

4.1 Stakeholder analysis

The following key stakeholder groups will be impacted by the development and implementation of the interactive animatronic creature.

P&P Projects: as a leading company in themed entertainment and the project partner, P&P Projects has a deep interest in the success of the animatronic creature. Their primary concerns include making sure the creature aligns with their design vision, engages visitors, and contributes to a positive guest experience within the theme park.

Myself (The Developer): as the designer and developer of the interaction engine, my primary interest lies in creating a robust, reliable, and innovative system that pushes the boundaries of animatronic technology. I am also committed to making sure the system is user-friendly, ethically sound, and meets the needs of all stakeholders.

Theme Park operators: the operators of the theme park where the animatronic creature will be deployed have a big interest in its performance and impact on guest satisfaction. Their concerns include the reliability and maintainability of the system, its safety for visitors, and its overall contribution to the park's entertainment value and revenue generation.

Theme Park visitors: ultimately, the success of the project hinges on the experiences of theme park visitors. Their interests include engaging in enjoyable and memorable interactions with the animatronic creature, feeling a sense of connection and wonder, and having a positive overall experience within the theme park.

4.2 Preliminary requirements

Together with P&P Projects a set of preliminary requirements has been established to guide the design and development of the interaction engine.

4.2.1 Functionality

Natural interactions: the interaction engine should enable the animatronic creature to engage in natural interactions with visitors. A Natural interaction is understood as one where the creature’s movement is smooth, transitions are not abrupt and do not contain unnatural jumps, and reaction times are within a determined time.

Multi-layered behavior: the system should incorporate a multi-layered architecture, combining deliberative, autonomous, and reactive behaviors to create the creature’s movement output.

Customization: the engine should allow for customization of the creature's behaviors and responses to fit specific theme park environments, storylines, or character personalities.

4.2.2 Usability

User-Friendly interface: the system could include a user-friendly interface for theme park operators to control and monitor the interaction engine and adjust parameters.

Reliability and robustness: the engine should be designed for reliability leading to consistent performance. In this way it minimizes the risk of technical failures or unexpected behaviors.

4.2.3 Ethical considerations

Privacy: the system should minimize the storage of personal data from visitors.

Safety: the system should not include any behaviors that can be perceived as threatening.

Inclusivity: the interaction engine should be designed to accommodate a diverse range of visitors, making sure that the creature is able to detect and interact with people of all ages, abilities, and cultural backgrounds.

4.2.4 Final table of requirements

For the requirements above, a final table was made including the must, should, could, will not have (MoSCoW) method for each requirement. This table was the final requirement list that guided the project.

Table 1 Requirement Analysis

Requirement	MoScoW
Natural interactions	Must have
Multi-layered behavior	Should have
Customization	Should have
User-Friendly interface	Could have
Reliability and robustness	Should have
Privacy	Must have
Safety	Must have
Inclusivity	Must have

4.3 Ideation of concepts

Two potential architectural concepts have been made for the interaction engine, each offering different approaches to processing information and guiding the animatronic creature's behavior:

4.3.1 Concept 1: multilayered rule-based system

In this concept, the decision-making process will be driven by a rule-based system, providing a structured framework for mapping perceived inputs to corresponding actions and responses. This system will operate on a multi-layered architecture, incorporating deliberative, autonomous, and reactive behaviors to create a dynamic and engaging character. Predefined rules will dictate the creature's responses based on combinations of sensor data, internal states, and contextual factors. Autonomous behaviors, such as breathing and blinking, will run continuously in the background to enhance lifelikeness, while reactive behaviors will be triggered by unexpected events to add spontaneity and realism. The system will prioritize and blend these different behavioral layers to achieve a natural and seamless flow of interaction.

4.3.2 Concept 2: hybrid system with AI integration

This concept integrates artificial intelligence (AI) into the decision-making process, alongside a rule-based system, to achieve an adaptive behavior. AI algorithms, trained on data collected from visitor interactions and environmental factors, would learn to recognize complex patterns, and generate responses that are more personalized and contextually aware. For example, the AI component could analyze a visitor's body language, tone of voice, and past interactions to infer their emotional state and intentions, allowing the creature to respond in a more empathetic and engaging way.

4.4 Preferred concept

While both concepts offer viable approaches, the centralized rule-based system (concept 1) is chosen as the preferred architecture for the interaction engine. The decision to utilize a rule-based system with layered behaviors as the core of the interaction engine comes from a consideration of the project's goals, constraints, and the overall theme park environment. This approach offers several advantages that align with the needs of both the developers and the stakeholders which will be described in the next paragraphs.

4.4.1 Consistency

Theme parks thrive on creating predictable and consistent experiences for their visitors. The rule-based system makes sure that the animatronic creature's behavior remains within expected boundaries, respecting the established character intent and providing a controlled experience for guests.

4.4.2 Explainability

The transparency offered by a rule-based system is important for theme park operators. The clearly defined rules provide a direct understanding of the decision-making process, making it easy to identify the reasons behind the creature's actions and making adjustments easy.

4.4.3 Ease of implementation

Given the time constraints and the complexity of the project, going for a rule-based system offers a practical advantage in terms of implementation. Developing and refining a comprehensive rule set is a more manageable task compared to the challenges of training and integrating AI models. This ease of implementation is particularly important for keeping the project within budget and time constraints.

4.5 Limitations of the preferred concept

4.5.1 Limited adaptability

Unlike AI-based systems, the rule-based approach may not adapt as readily to unexpected situations.

4.5.2 Rigidity

An overly rigid rule-based system may result in repetitive or robotic-feeling interactions. To avoid this, the system can incorporate elements of randomness and variation within the defined rules.

This page has been intentionally left blank.

5. Realization

During the realization phase, an interaction engine was made which fulfilled the requirements set in the specification phase. A general description of the engine can be found below:

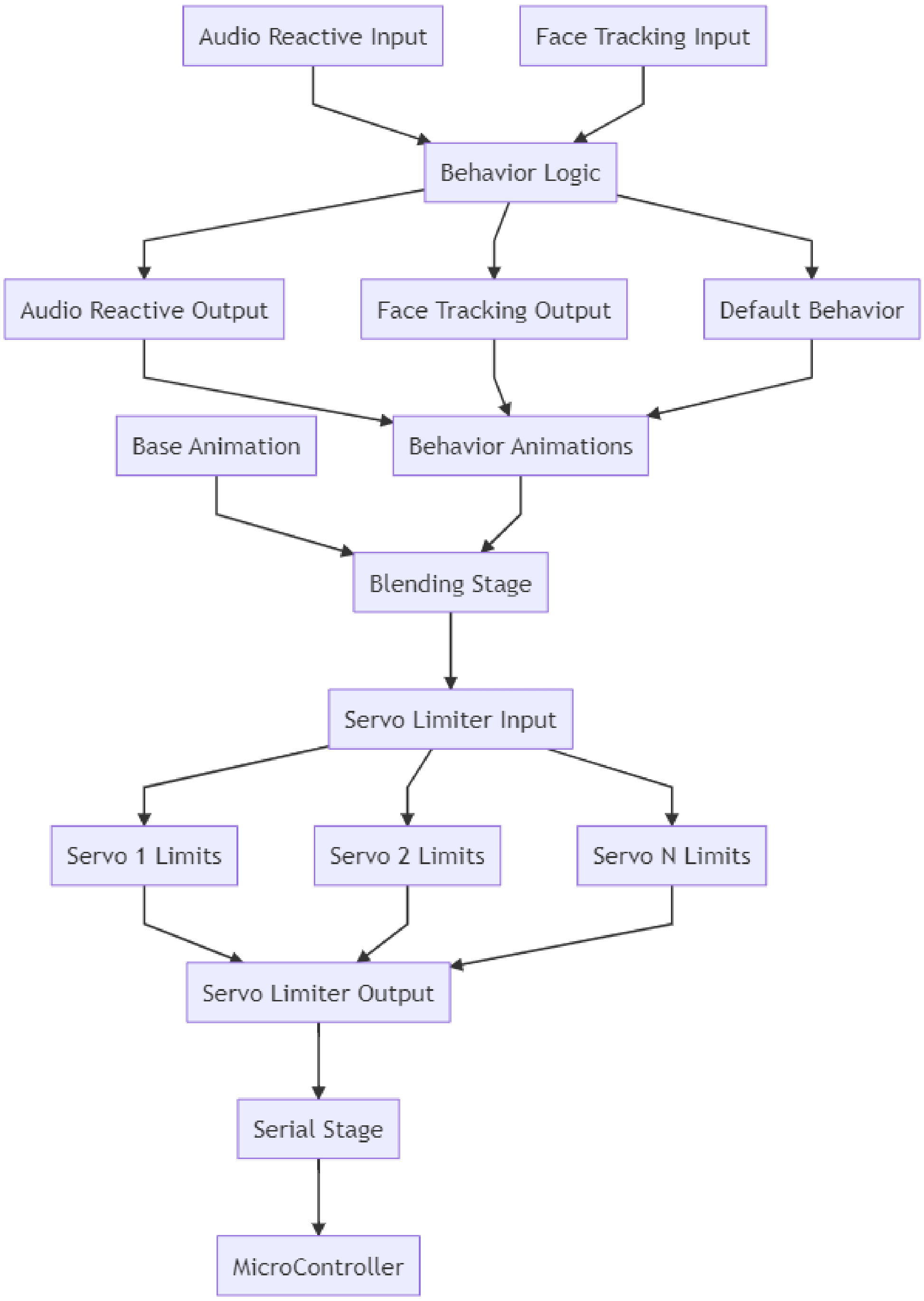


Figure 7 Interaction engine description.

5.1 High level interaction engine view

To understand the interaction engine's function, let's imagine it as a series of interconnected stages that blend and sequence animations, incorporating live inputs to create a dynamic and believable character. The engine can be visualized as a pipeline, starting with the creation of a base animation, adding behavioral animations, and finally transmitting position commands to the animatronic creature's servos.

The interaction engine consists of a series of blended and sequenced behaviors mixed with live inputs.

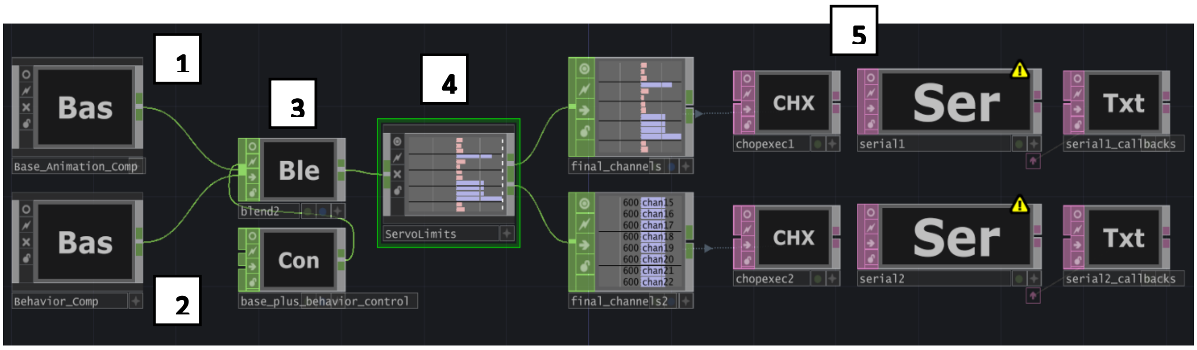


Figure 8 High Level view of the interaction engine

All behaviors and animations in the interaction engine are represented in servo position vs time. In a high level, it consists of a base animation [Figure 8, 1] which is blended with the behavior animations [Figure 8, 2] in the blending stage [Figure 8, 3]. This is then fed to a servo limiter [Figure 8, 4] and finally sent over serial to two different controllers in the serial stage [Figure 8, 5]. Each of these elements have their own internal functioning, but the focus will be on the general overview first.

The base animation consists of an animation including breathing, slight movements and twitches which are then blended with the behaviors coming from the behavior composition. The amount of each animation is then decided in the final blending [Figure 8, 3]. This is percentage based. In general, through experimenting and playing with the interaction engine it has been found that a percentage of 20% base animation and 80% behavior animation yields a good balance where the base animation is faint but perceivable, while the behavior composition is predominant. This percentage is adjustable in the blending stage to achieve the best results. This blending uses an additive blending technique, which is found in all blending that happens in the system. A description of how this blending works can be found in section 5.4 “additive blending”.

The blended animation coming from the final blending stage then is fed to the servo limit setting stage [Figure 8, 4]. This stage sets the limits of the physical mechanisms connected to the servos. It maps the position values coming from the previous stages which are normalized (0 to 1) to a specific range of pulse with modulation for each servo (for example, 600 to 2500). The pulse with modulation is directly proportional to the servos position. This step is matched to the 21 servos currently in the animatronic character. They have been set to correspond to their mechanical limits to not exceed the physical limits of movement for the servos. This is also adjustable; in case a mechanism is modified, or this system is to be integrated to a different character.

Finally, the mapped channels are split into two serial communication channels for the two different controllers that the character’s servos are connected to. In the serial stage, a port can be selected where the commands can be sent. The commands are sent with the format Servo Number, Value \n. These commands are then received by the controller which triggers that movement for the servo. For example, if the value sent is 1,600\n then the servo 1 gets sent 600PWM which moves it to its 0 position. The position mapping is based on a servo's range of 600PWM as a minimum (0 position) and 2500PWM as a maximum (180 degrees in this servo).

5.2 Base animation

The base animation serves as the foundation for the creature's movements and expressions. It comprises a blend of subtle, rhythmic movements, including breathing consisting of a gentle, undulating motion of the head and shoulders to simulate breathing. And almost imperceptible muscle twitches to add a touch of realism and suggest a sense of life.

This base animation [Figure 9, 1] is pre-animated and has a looping length of 40 seconds. This animation length is the same for all animations in the system. The base animation also has a scaling component [Figure 9, 2]. This scaling component serves as added adjustability in the system. Its purpose is to tone down and up the base animation, to modify its effect in the final motion. Finally, the result of this animation is sent to the final blending stage described in 5.1 [Figure 8, 3].

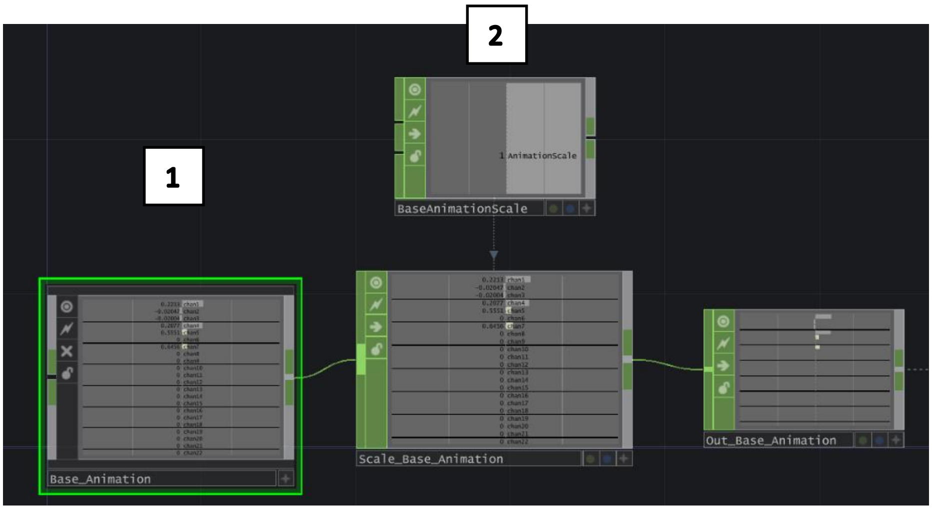


Figure 9 Base animation

The base animation can be animated using key points [Figure 10, 1]. Each key frame has also added adjustability using Bezier curve modifiers. The animation portrayed in [Figure 10] is an example of the breathing motion present all the time in the character.

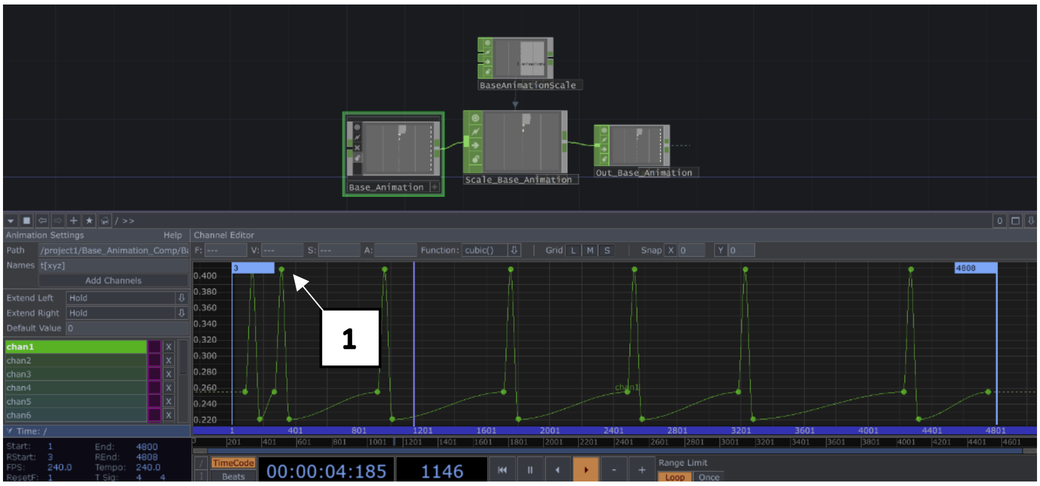


Figure 10 Base animation programming.

5.3 Behavior composition

The behavior composition is the most complex element of the system, as it houses all the elements that trigger and blend between animations and live inputs. First a high-level description of the behavior composition will be provided, and more detailed components explained further on.

5.3.1 High level behavior composition

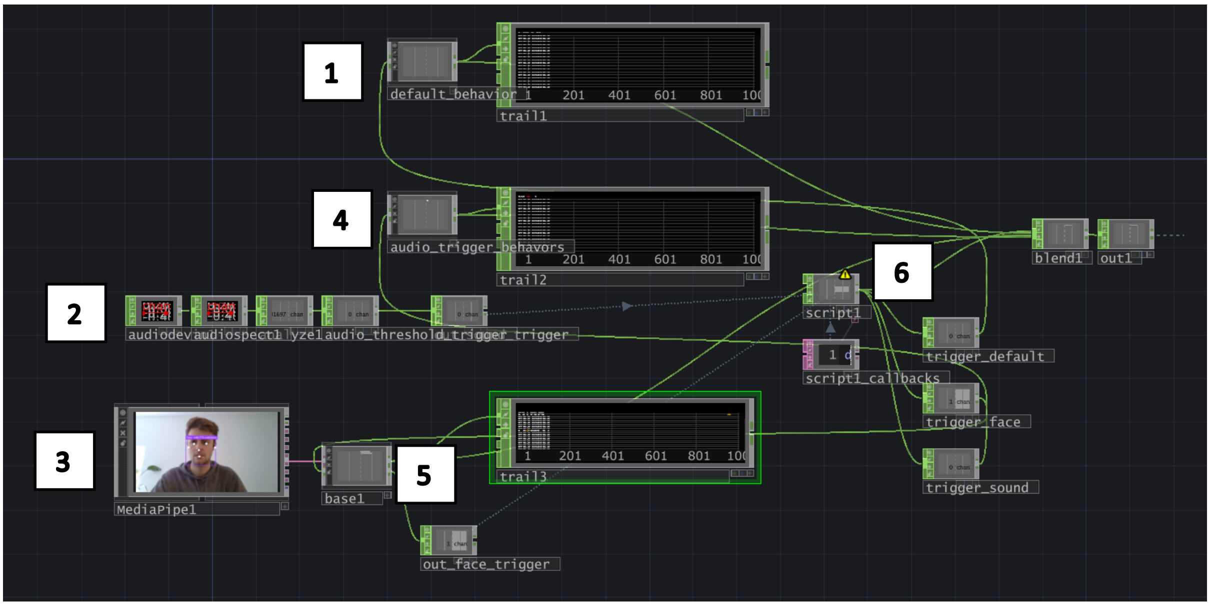


Figure 11 High level Behavior composition.

The Behavior composition is divided into 6 distinct parts: the default behavior [Figure 11, 1], the audio reactive input [Figure 11, 2], the audio reactive output behavior [Figure 11, 4], the face tracking input [Figure 11,3], the facial tracking output behavior [Figure 11, 5] and the behavior logic [Figure 11, 6].

The default behavior [Figure 11, 1] is a collection of animations which are blended and randomized when triggered, this is a standard element which has been used several times in the interaction engine It will be referred to as the “Standard animation bank”. It will be further explained in section 5.3.1. This is the default state of the behavior composition, meaning that if no other behavior is triggered (no faces on sight and no audio reactiveness) then the character will transition into this state.

The audio reactive input behavior [Figure 11, 2] takes as an input an audio signal from the microphone, and through a series of filters and limits sets a trigger to true when a loud sound is detected. This trigger is later assessed by the behavior logic.

The audio reactive output behavior [Figure 11, 4] is another instance from the standard animation bank. When triggered, a random animation starts playing.

The facial tracking input behavior [Figure 11,3] takes as an input a video feed from the camera and adds facial tracking points to the detected faces. It can detect and pass up to 10 faces. When a face is detected, it sets a trigger to true that is later assessed by the behavior logic.

The facial tracking output behavior [Figure 11, 5] is a slightly modified version of the standard animation bank, it has an extra layer where the coordinates of the faces are blended to the output channels. This makes sure that the characters follow people around. When triggered, a certain animation starts playing with the facial coordinates mixed in.

The behavior logic [Figure 11, 6], takes decisions for when to blend in and out the default, audio and facial tracking behaviors. It does this based on the audio and facial triggers coming from the audio and facial inputs. The result of this behavior logic is a single behavior animation which is then passed to the general blending stage in 5.1, [Figure 8, 3].

5.3.2 Standard animation bank

The standard animation bank [Figure 12] is an element which is repeated several times within the interaction engine. This animation bank has a trigger input [Figure 12, 1] and an animation output [Figure 12, 2]. When triggered, it will cycle between the 10 animations contained in it randomly [Figure 12, 3]. When transitioning between two animations, it blends both animations during a certain amount of time before it continues to the following animation. It does this based on a series of variables and an additive blending method contained in a python script [Figure 12, 4]

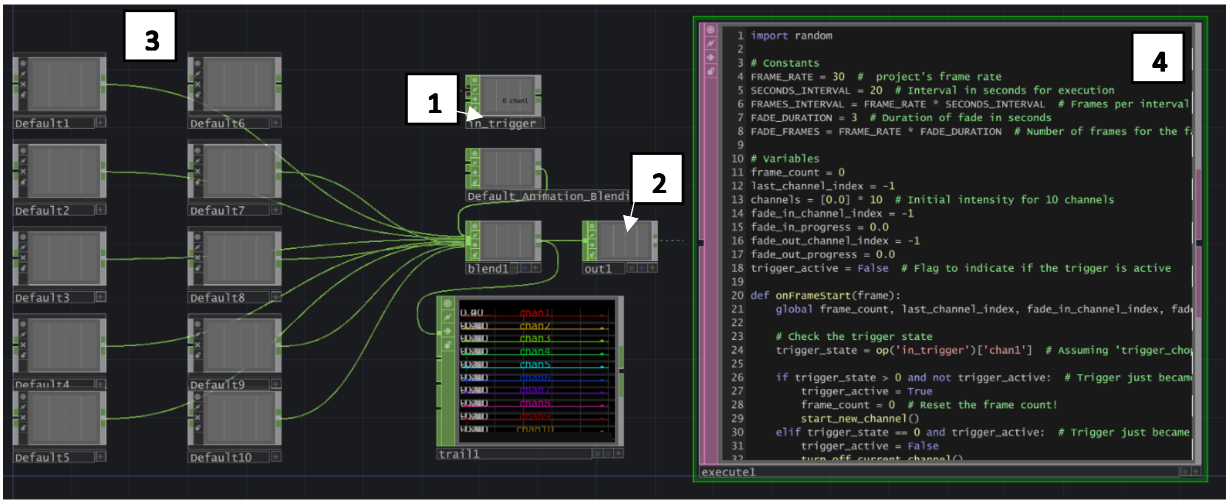


Figure 12 Standard animation bank.

As seen in [Figure 13] The animations contained within the standard animation bank can be animated in the same way described in chapter 5.2 base animation. All animations in the interaction engine are compatible, so they can be copied and interchanged as necessary.

Figure 13 Standard animation bank animating example.

5.3.3 Standard animation bank with facial tracking

To achieve the desired effect with the facial tracking, an intermediate layer had to be added to the standard animation bank [Figure 14]. This layer takes care of extracting the center positions of the faces and blending it into the standard animation bank used in the facial tracking.

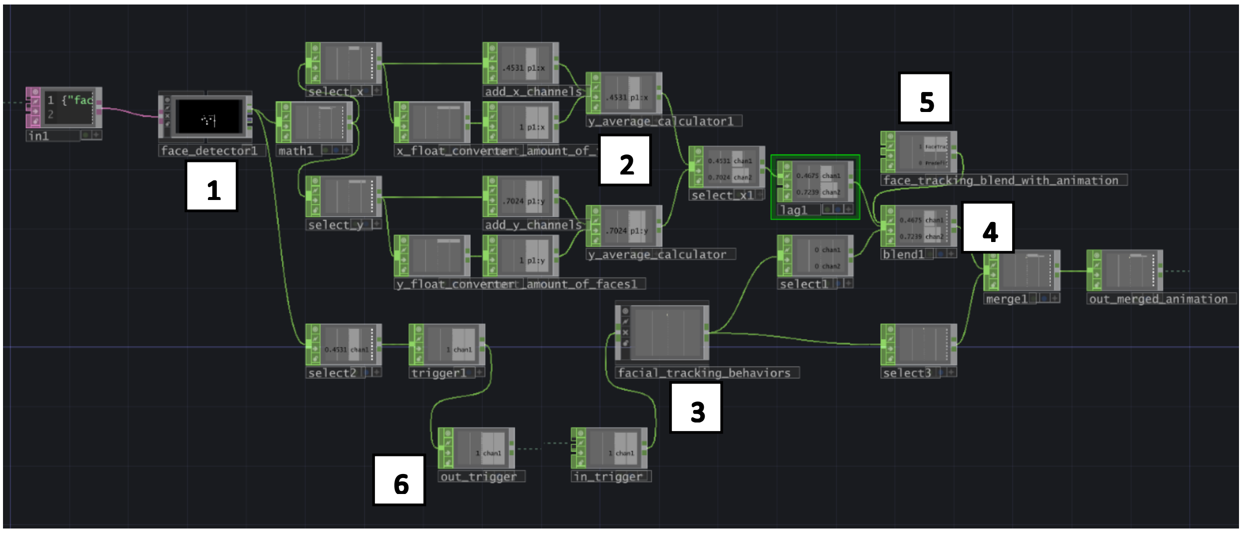


Figure 14 Facial tracking composition.

This composition layer has as an input the coordinates extracted from the facial tracking composition [Figure 14, 1]. Which is a standard OpenCV touch designer plugin: Mediapipe [27]. These variables are then split and mapped [Figure 14, 2], to be blended into the animation coming from the facial tracking standard animation bank [Figure 13, 3]. This approach also uses additive blending [Figure 14, 4], which keeps the motion animated and adds the facial tracking variables. It is possible to adjust the percentage of participation of each [Figure 14, 5], to achieve the desired facial tracking without losing the background motion. The percentage of participation which has given the best results is 20% for the animation and 80% for the facial tracking variables. The variables are mapped in a linear way, for example, the position of the face in the camera measured in pixels (0 to 1080) corresponds to the rotation of the neck in degrees (-30 to 30). This is perhaps not the most accurate mapping, but for this application it works well.

Finally, this composition layer has a trigger [Figure 14, 1] which is activated when there is a face present, which is later transferred to the behavior logic as specified in 5.3.1 [Figure 11, 6].

The animations in this standard animation bank are also compatible with the rest animations of the system, making it possible to interchange when necessary.

5.3.4 Behavior logic

The roles of the default animation, sound trigger behavior and facial tracking behaviors have been explained. In this section, the method for blending and selecting between them is explained. This is accomplished by the behavior logic [Figure 15], it fades and blends between the three behaviors based on a set of priorities and variables.

The default logic lives in a script which implements the behavior blending system [Figure 15, 1], transitioning between three distinct animation sources: default, sound, and face. The script uses a Constant [Figure 15, 2] as the central hub for controlling animation blends, with each channel representing a behavior and its value (0.0 to 1.0) dictating the blend amount. Triggers for initiating sound and face behavior behaviors are provided by the inputs of both respectively. The default animation serves as the fallback, playing continuously unless overridden.

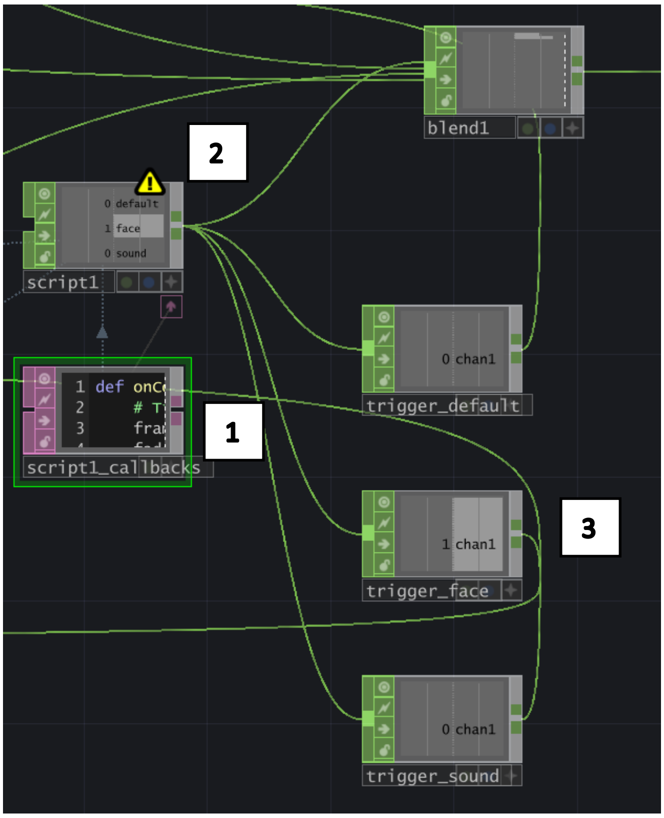


Figure 15 Behavior state Logic

To introduce a hierarchy in animation selection, a priority system is implemented using priority constants with values 0, 1, and 2 denoting ascending priority. In the event of concurrent triggers, the behavior with the highest priority is blended in.

Moreover, the face behavior has a time-based control mechanism. A maximum on-time limits its continuous play duration. After exceeding this threshold, the face behavior source is forced to deactivate, to keep the behavior diversity. Also, a re-trigger delay, prevents immediate reactivation of the face behavior source after deactivation, this is done to have a more dynamic flow.

The script also has output triggers [Figure 15, 3] which are set to true when a given behavior is active. This is then used to make sure that when a behavior is starting to be blended in, it is in its starting state.

Finally, the script uses a frame-based approach for smooth transitions between behaviors. Fade durations are defined for each behavior, allowing for gradual increases and decreases in blend values over time.

This concludes the Behavior composition explanation, its output will then be blended with the base animation and transformed and scaled to be sent to the servos.

5.4 Additive blending

A method that is used in all blending instances in the system is additive blending [Figure 16]. This technique makes sure that all the input animations are considered based on a percentage value to produce the output animation. The most important aspect of this blending technique is that the percentage value of animation must be faded in and out for the animations to be blended properly. If the channel is directly turned off, there is a jump in the animation that does not look natural.

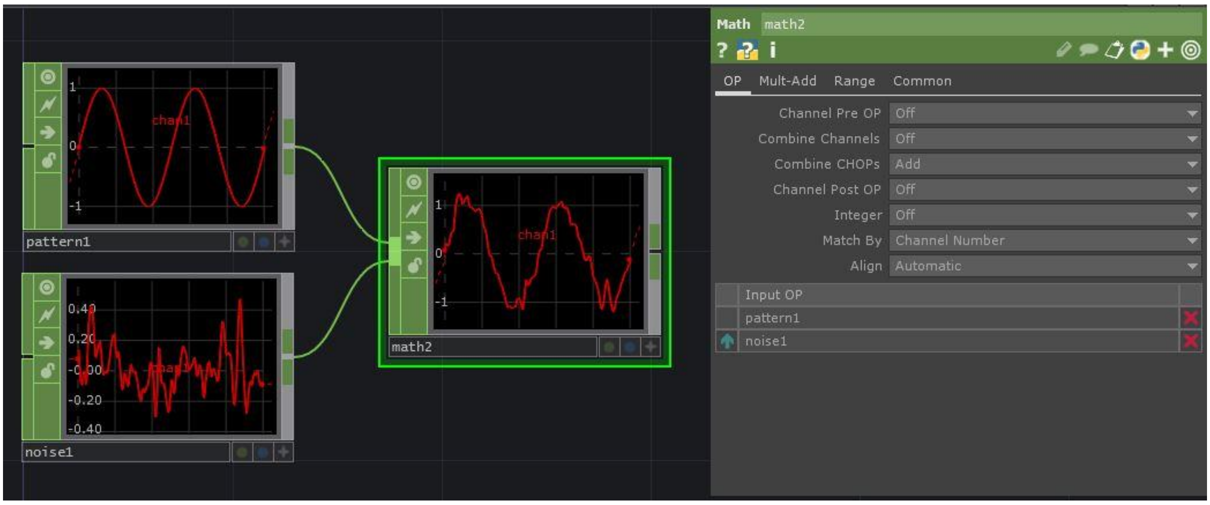


Figure 16 Additive Blending.

5.5 Servo limiting

To limit the motion of the servos to the limits of their mechanism, the servo limiting composition was implemented. This composition makes sure that the servos always operate within their acceptable range of positions, without losing resolution by using standard limiters. This is done individually to all servos, as can be seen in [Figure 17].

Figure 17 Servo limiting.

In a more detailed view [Figure 18], it can be seen how each channel corresponds to one servo, and the limits for it are individually set. This is done by mapping the value coming from the behavior which is ranges between 0 to 1 [Figure 18, 1] to the range of the mechanical boundaries of the servo. The minimum and maximum values for this are 600 to 2500 [Figure 18, 2], but depending on the mechanism that is moved by the servo this range could be reduced.

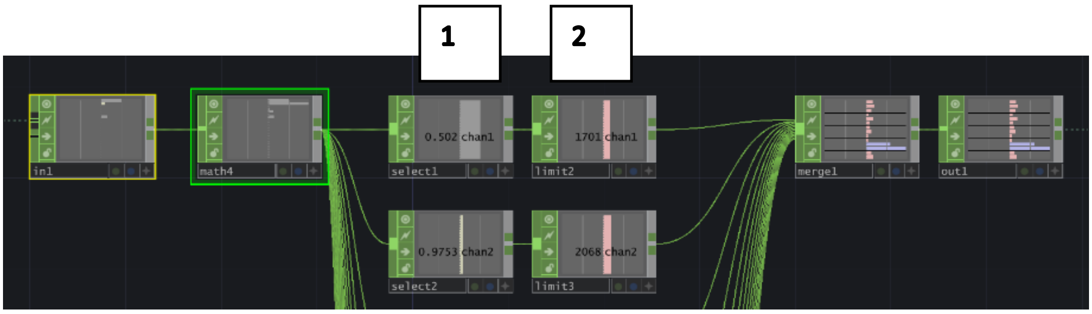
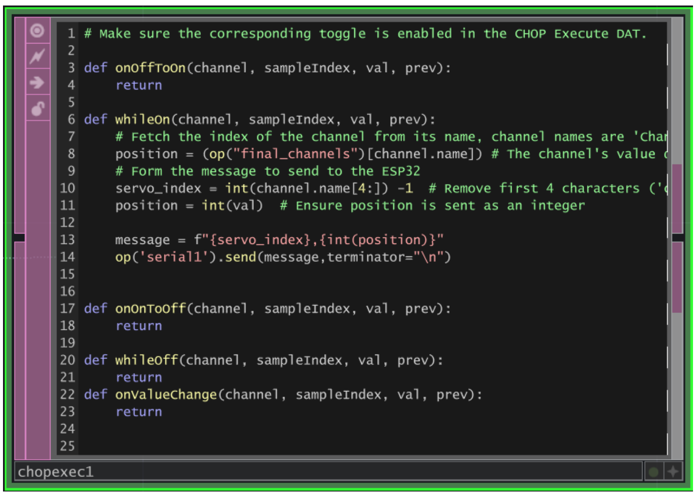


Figure 18 Detailed view servo limiting.

5.6 Serial communication

The animation channels coming from the servo limiting are sent over serial to the servo controllers. The servo controllers are expecting the positions of the servos in the format of Channel Number, Position/n. To accomplish this, a script has been written that takes the channel from the final channels CHOP and sends it over serial using this protocol. This script [Figure 19] Takes the channel name and removes the unnecessary information, to then send it using the expected protocol.



```
# Make sure the corresponding toggle is enabled in the CHOP Execute DAT.

def onOffToOn(channel, sampleIndex, val, prev):
    return

def whileOn(channel, sampleIndex, val, prev):
    # Fetch the index of the channel from its name, channel names are 'Cha
    position = (op("final_channels")[channel.name]) # The channel's value
    # Form the message to send to the ESP32
    servo_index = int(channel.name[4:]) -1  # Remove first 4 characters ('
    position = int(val)  # Ensure position is sent as an integer

    message = f"{servo_index},{int(position)}"
    op('serial1').send(message,terminator="\n")


def onOnToOff(channel, sampleIndex, val, prev):
    return

def whileOff(channel, sampleIndex, val, prev):
    return
def onValueChange(channel, sampleIndex, val, prev):
    return


```


Figure 19 Serial Script

This process is done twice, as the servos are connected to two different controllers. Each of the scripts sends the serial command to a different port. The controllers used for this purpose are two ESP32WroomDA module. They were selected because of their speed and ease of connection. The code in the microcontroller takes the serial commands and directs It to the corresponding servo. This code can be found in Appendix A.

5.7 Hardware setup

To receive the commands from the interaction engine and transform them into PWM sent to the servo motors, a pre-existing hardware setup was reused. This hardware setup consists of a PC with the "Touch Designer Perform" program running the interaction engine. This PC also receives the feed from the webcam for facial recognition, and the feed from the microphone. Both ESP32WroomDA modules are connected to this PC, and their selected output pins are then connected to the servo motors. A detailed overview of this hardware can be found below:

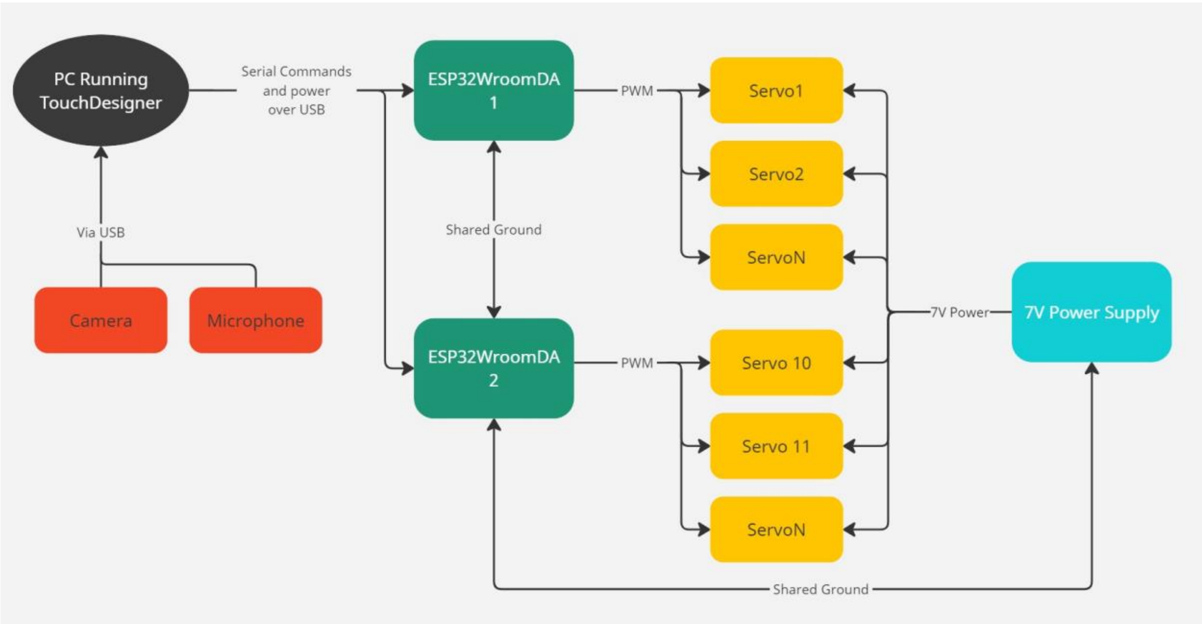


Figure 20 Hardware Diagram

5.8 Filling the animations.

5.8.1 Brainstorming

To fill the animations inside the interaction engine, an ideation session together with the animation experts at P&P projects took place on the 28th of May 2024. For this activity, ethical approval by the examination board was granted with a positive advice, with the application No. 240104. The information letter and consent forms can be found in the Appendix B.

The session started by envisioning the creature's default behaviors, those subtle actions that would convince onlookers it possessed a life of its own even when at rest. The animators, drawing upon their expertise in animal behavior and character animation, suggested a range of actions: a curious head tilt, a gentle shake of the head as if dispelling a pesky thought, even a playful swat at an imagined insect.

The discussion then moved towards reactive animations, brainstorming how the creature might respond to a sudden noise, a bright light, or the appearance of a visitor. Ideas ranged from a startled jump and wide-eyed surprise to a curious head turn and a playful attempt to engage with a passing hand. By the end of the meeting, we had compiled a list of potential animations designed to fill the creature with the illusion of life.

The full list of behaviors derived from this meeting can be found on Appendix C.

Figure 21 Discussing behaviors with the animators of P&P Projects.

5.8.2 Animating the character.

After the brainstorming had taken place, the animations selected were animated into the interaction engine. A lot of time and consideration went into creating the routines that would highlight the creative intent of the character. These animations were based on the generated animation list, but also in the responsive behavior that we aimed to introduce in the creature.

Behaviors such as the facial tracking behavior and the base behavior rely more on specific movements of the creature to convey their purpose. A list of those movements can be found in the following page.

Table 2 Creature movements

BEHAVIOR	MOVEMENTS
BASE BEHAVIOR	Head up and down
	Head tilt
	Right ear up and down
	Left ear up and down
	Mouth side to side
	Mouth up and down
	Eyelids open and close
DEFAULT BEHAVIOR	Left foot side to side
	Left foot up and down
	Right foot side to side
	Right foot up and down
	Neck side to side
	Neck up and down
	Neck tilt
	Nose up and down
	Hands side to side
	Hands front and back
	Mouth side to side
	Mouth up and down
	Eyelids open and close
	Eyes side to side
	Eyes up and down
FACIAL TRACKING BEHAVIOR	Eyes side to side
	Eyes up and down
	Head side to side
	Head up and down
	Head tilt
	Eyelids open and close
	Mouth
AUDIO REACTIVE BEHAVIOR	Left foot side to side
	Left foot up and down
	Right foot side to side
	Right foot up and down
	Neck side to side
	Nose up and down
	Hands side to side
	Hands front and back
	Mouth side to side
	Mouth up and down
	Eyelids open and close

As it can be seen in this table, the behaviors rely on specific movements that match the creative intent of the behavior. However, the behavior animation itself also includes all other movements, but they are not active or are not prominent enough to include them as part of that behavior.

6. Results

This section dives into the interaction engine's capabilities, illustrating its functionality through practical examples. We'll analyze the interplay between the base behavior, default behavior, and how these integrate with the facial tracking and audio reactive behaviors.

6.1 Base behavior

The base animation, the heart of the creature's lifelike presence, operates continuously in the background. It infuses subtle movements into the character, preventing a static or "lifeless" appearance when no other behaviors are active.

Let's break down the base behavior, illustrated in [Figure 22]:

Figure 22 Base Behavior Breakdown

[Figure 22 , 1] This snapshot captures the creature at the peak of its inhale. The head tilts slightly back, while the hands rise, mimicking a natural breathing rhythm.

[Figure 22 , 2] The midpoint of the breathing cycle showcases the head returning to a neutral position. The shoulders simultaneously lower, creating a smooth transition between the inhale and exhale.

[Figure 22 , 3] Captured at the peak of exhalation, the image displays the head tilting slightly forward. The shoulders lower further, completing the exhale motion. This subtle contraction adds a layer of realism to the breathing cycle.

It is important to consider that the movement seen in [Figure 22] is exaggerated and going to its limits. However, when it is blended with the behavior composition its reduced to 20%. The length of this cycle is approximately 4 seconds.

This constant movement makes the creature appear lifelike, even when at rest, creating a foundation for layering in more complex behaviors.

6.2 Default behavior

The Default Behavior represents a step up in complexity from the always-present Base Behavior. Rather than subtle, constant movements, the Default Behavior has a series of longer, more deliberate actions that showcase a range of lifelike mannerisms. These 40-second animations cycle randomly when no external stimuli (audio or facial) are present, creating the illusion that the creature is autonomous and curious about its surroundings.

Here is an example for one such animation within the Default Behavior set to illustrate this:

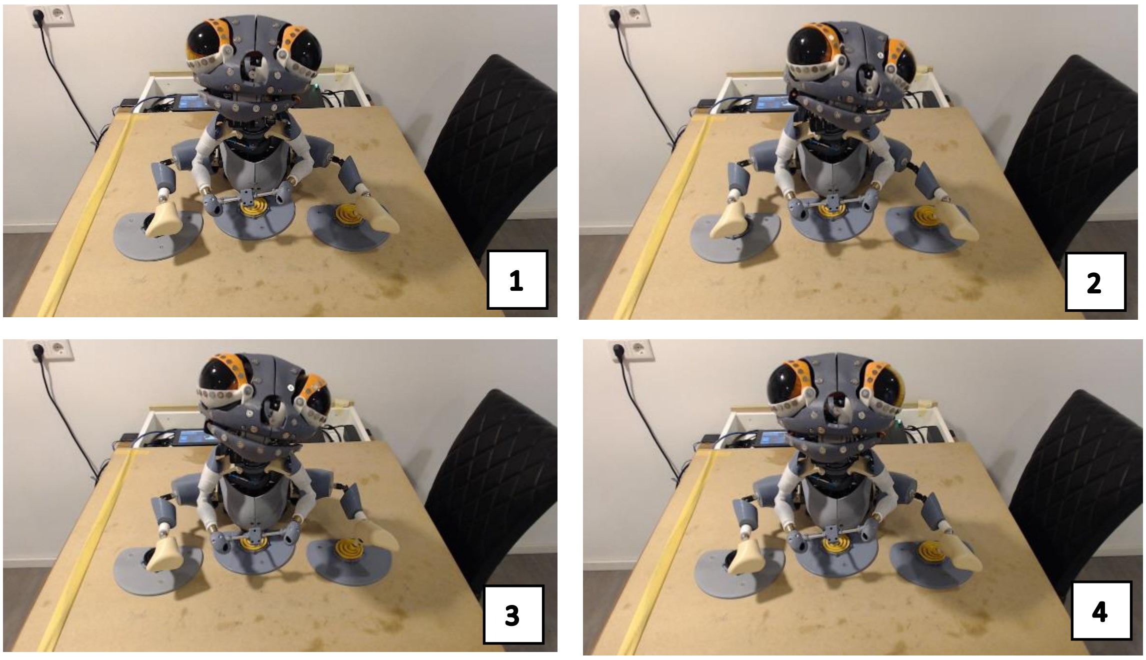


Figure 23 Default Behavior: "The Curious Glance"

[Figure 23, 1] The animation begins with the creature in its neutral pose. This provides a moment of stillness, drawing attention to the start of the upcoming action sequence.

[Figure 23, 2] Driven by curiosity, the creature's head slowly begins to turn to the left. The movement is fluid and natural, as if something has piqued its interest in the environment.

[Figure 23, 3] The gaze fixates on a point to the left, held for a few seconds, suggesting a focused observation of something unseen. This pause adds a layer of believability, mimicking how a living creature would focus its attention.

[Figure 23, 4] As if satisfied with its inspection, the creature's head gradually returns to its center position. This smooth transition prevents the movement from feeling abrupt or mechanical.

This example shows how the Default Behavior uses longer, more complex animation cycles to show lifelike actions and mannerisms. The Default Behavior creates an illusion of autonomy, making the creature appear thoughtful, observant, and engaged with its environment, even when alone.

6.3 Facial tracking behavior

The facial tracking behavior adds a new dimension to the creature's interactions, allowing it to acknowledge and engage with visitors. When a face is detected, the creature smoothly transitions from its default behavior to a set of animations designed to acknowledge and follow the individual.

[Figure 24] showcases this interaction:

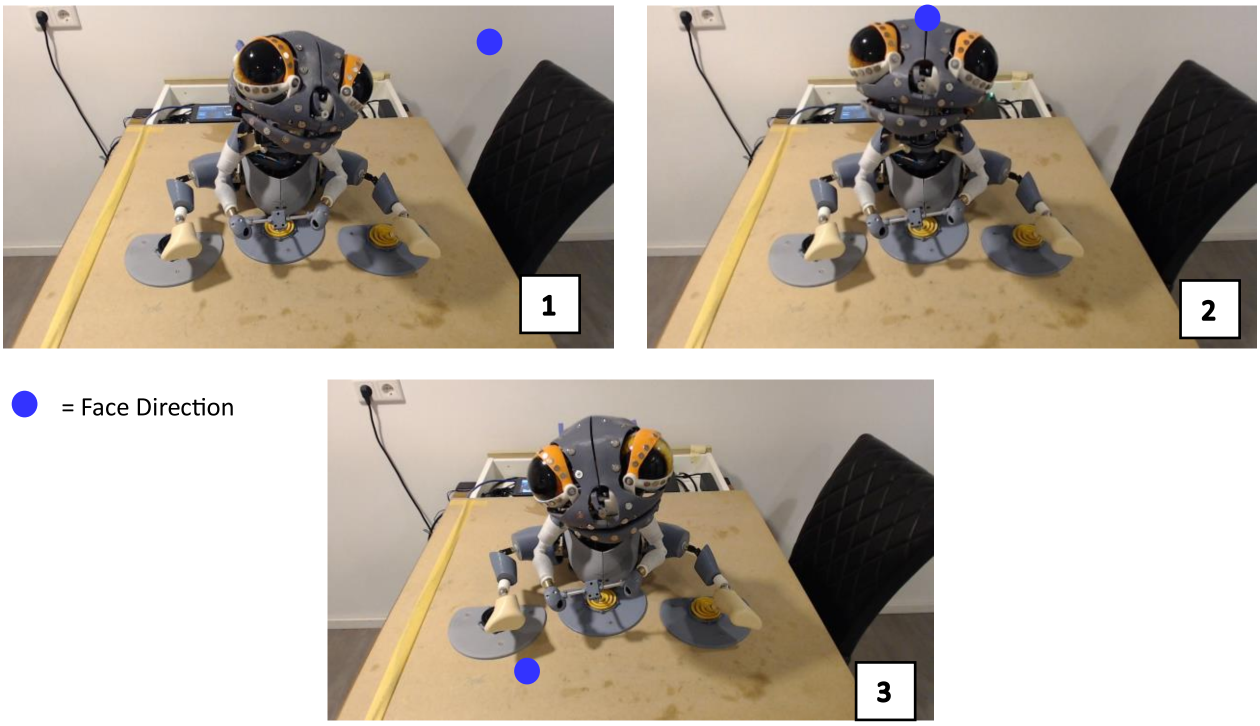


Figure 24 Facial Tracking Behavior

[Figure 24, 1] As a visitor approaches, the creature's gaze snaps towards the individual, demonstrating awareness and initiating engagement.

[Figure 24, 2] Maintaining eye contact, the creature's head smoothly turns to follow the visitor's movements, showcasing a natural response to their presence.

[Figure 24, 3] The creature's gaze and head position continuously adjust in sync with the visitor's movement, creating a feeling of being followed and acknowledged.

This direct engagement, driven by facial tracking, creates a personalized experience for the visitor, making them feel acknowledged and drawn into the interaction. When there are two visitors at the same time, the creature will either choose to look at both of them during 5 second intervals or look at the center point between them. As this creature has eyes which allow for looking at two different areas without making it feel strange as a human would.

6.4 Audio reactive behavior: responding to the environment

Complementing the other behavior layers, the audio reactive behavior allows the creature to respond to its auditory environment. When a loud sound is detected, the creature transitions from its default or facial tracking behavior into one of 10 designated set of audio reactive behaviors, mirroring a natural startle response.

Let's analyze one of these responses through [Figure 25]:

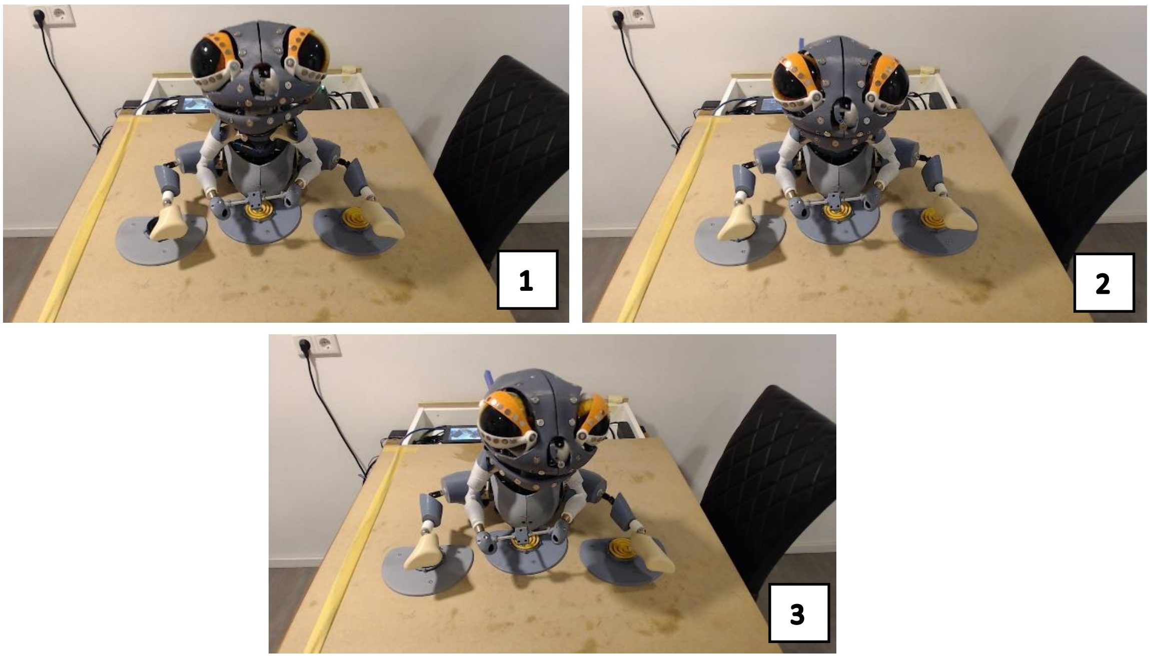


Figure 25 Audio Reactive Behavior: Startle and Return

[Figure 25, 1] Upon detecting a sudden loud noise, the creature's eyes widen, emulating a natural startle response. This immediate reaction contributes to the feeling that the creature is genuinely surprised.

[Figure 25, 2] The creature's head quickly retracts, further emphasizing the startle reflex and suggesting a degree of caution or apprehension in response to the unexpected sound. The eyes also adapt to show a careful consideration of the sound.

[Figure 25, 3] Once the sound is done, the creature gradually returns to its neutral position, its movements slow and deliberate, conveying a sense of regaining composure after the surprise. The eyelids are also semi-closed during this time.

It's important to keep in mind that this is only one of many possible reactions to an audio input, they are different but keep the same principle in mind. The responsiveness to audio cues reinforces the illusion of life, making the creature's reactions feel less pre-programmed and more in tune with its surroundings.

6.5 Blending behaviors

The true strength of the interaction engine lies in its ability to blend between these distinct behavior layers, creating a dynamic and believable character. The transition between behaviors is managed by the behavior logic, which prioritizes, and fades animations based on a combination of triggers and timers.

For instance, if a visitor is interacting with the creature, the facial tracking behavior takes precedence. However, if a loud noise occurs during this interaction, the audio reactive behavior momentarily overrides the facial tracking, mimicking a realistic shift in attention. Once the sound subsides, the creature smoothly transitions back to engaging with the visitor.

This fluid blending of behaviors ensures the creature's actions never feel disjointed or robotic. Instead, they flow naturally, responding appropriately to a range of stimuli and maintaining a captivating illusion of life.

This page has been intentionally left blank.

7. Evaluation

This chapter presents an evaluation of the tested multi-layered interaction engine, assessing its performance against the predefined requirements outlined in the earlier section. The evaluation methodology had three main parts: observation, metrics, and expert opinions.

7.1 Natural interactions (MUST HAVE)

The goal of natural interactions is the core of this project. To create an animatronic character that goes beyond pre-programmed routines, capable of responding to its environment and visitors in a way that feels organic and engaging.

7.1.1 Observational analysis:

During two independent observation sessions taken place on the 12th and 13th of June 2024 at P&P Project's location, the creature's behavior was analyzed. Particular attention was given to the smoothness of transitions, responsiveness to stimuli and the realism of the animations. The team present was formed by a production expert with background in manufacturing for theme parks, a project manager with field experience in top-of-the-line theme parks, an art director with over 15 years of experience working in theme parks, and myself the developer. For this activity, ethical approval by the examination board was granted with a positive advice, with the application No. 240104. The information letter and consent forms can be found in the Appendix B.

Regarding the smoothness of transitions, the blending between the base, default, audio reactive, and facial tracking behaviors were assessed for abruptness and unnatural jumps. The goal was to achieve seamless transitions that mimicked the fluidity of natural movement. The creature's reaction time to both audio cues and the appearance of faces within its field of view was monitored. A fast, but not too reactive response was important to maintain a sense of believability.

The individual animations comprising the default and reactive behaviors were evaluated for their lifelike qualities. Did the movements appear deliberate? Did they convey a sense of curiosity or surprise when appropriate?

7.1.2 Metrics:

Reaction Time (Audio)

A test was made to assess the speed and reaction time of the creature towards an audio input. This was done by triggering the audio detection input with a loud noise and assessing the results across the output graphs. To simplify the analysis of the results, one channel was isolated for this purpose. In [Figure 26] the delay for tripping the audio trigger can be seen. The experiment was repeated 10 times, the mean trigger delay resulted being 1.24 seconds.

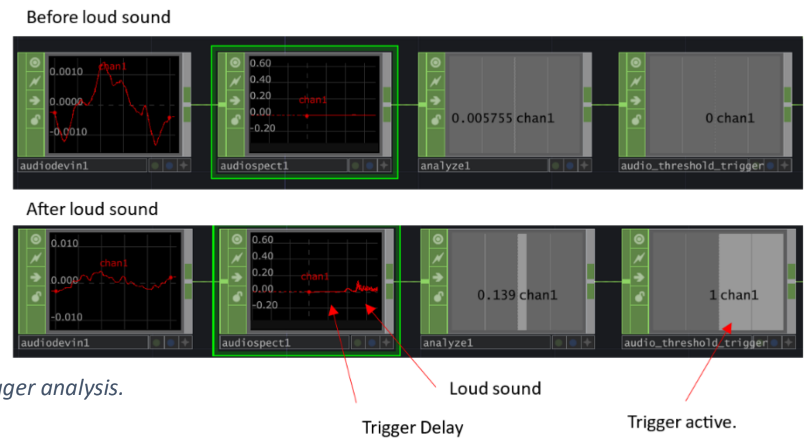


Figure 26 Sound trigger analysis.

However, the trigger delay is not the only variable at play while starting audio behaviors. This trigger delay starts a chain reaction which makes the sound reactive behavior visible. Namely the fading delays of the animation blending come at play. These are intentional delays which make sure that the animation blends smoothly with the underlying behavior. In [Figure 27] this can be exemplified:

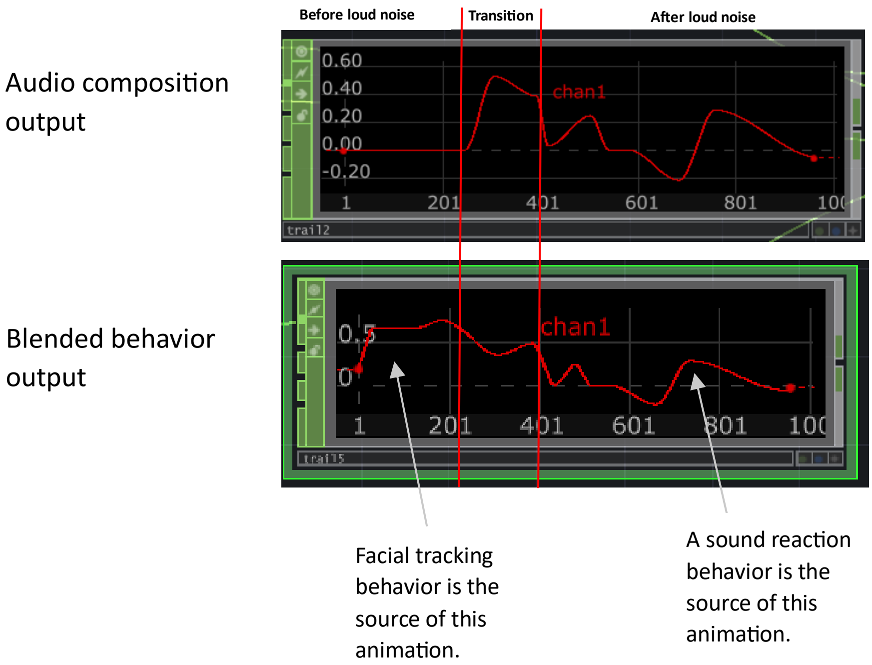


Figure 27 Sound reaction blending.

The total delay between when the sound is triggered and when the behavior is solely coming from the sound behavior is 4 seconds. This is an adjustable parameter, and 4 seconds has worked very nicely for this application. This makes the total delay between the loud noise and when the reactive behavior is perceptible 2.24 seconds considering the behavior being perceptible when it's at 25% of its blending.

Reaction Time (Facial)

A test was made to assess the reaction time between when the creature "sees" a face and when it reacts to it was also measured. Starting with the detection delay, which was measured by exposing a

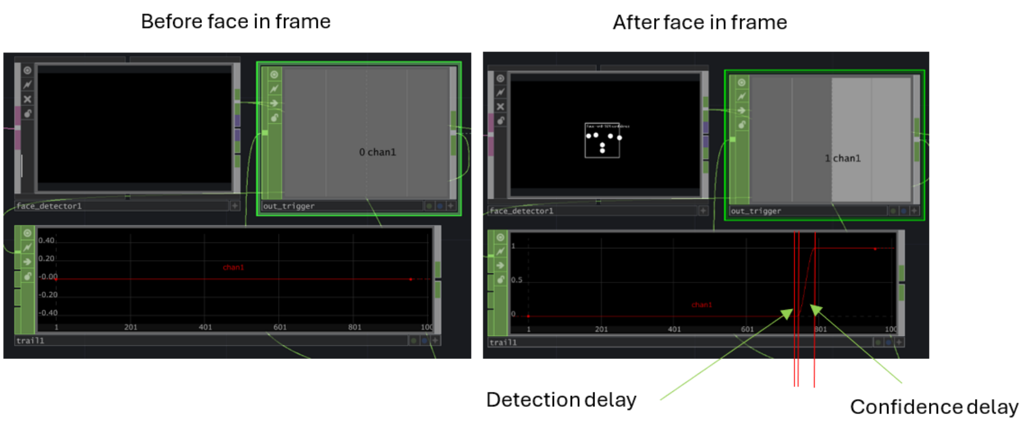


Figure 28 Facial reaction time.

face to the camera and measuring the time it took for the facial recognition algorithm to provide variables for that face. This test was repeated 10 times for accuracy, and it can be seen in [Figure 28]:

The detection delay had a mean of .36 seconds. However, to the detection delay we must add the confidence period. This period is an arbitrary time to get rid of false positive face detections. This period waits for the face to be present in frame for at least one second before the rest of the reactions are triggered. Therefore, the delay between when a face appears in frame and when the face tracking behavior is 1.36 seconds.

In the same manner as the audio reactive behavior, how fast the facial tracking is noticeable also is dependent on the blending with other animations. These are international delays meant to make sure that the facial tracking behavior is blended properly with underlying behaviors. This can be seen in [Figure 29].

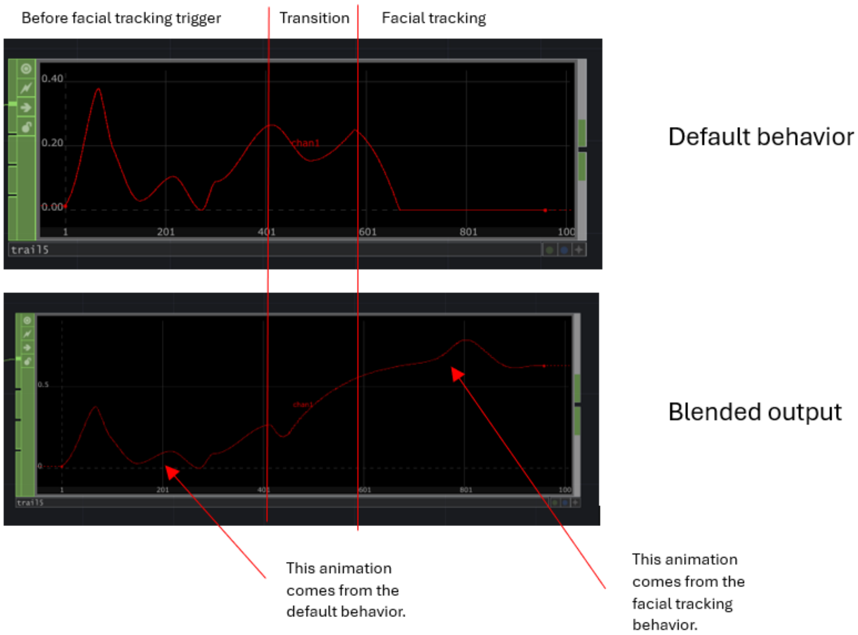


Figure 29 Facial tracking behavior.

As can be seen in [Figure 29], this example starts with the default behavior being the background behavior. When the facial tracking Is triggered, a transition period starts where the default behavior and the facial tracking are blended and faded together, this is an intentional delay of 4 seconds. Finally, only the facial tracking behavior remains.

In total, the time between when a face shows up in front of the character, and when we can notice a reaction there are 2.36 seconds, considering the behavior to be noticeable when its at 25% of fade in.

7.1.3 Findings in this evaluation

Observations revealed that the interaction engine was good in creating a sense of natural interaction. The transitions between behaviors were consistently smooth, with no big shifts in movement or posture. The creature reacted to both audio and facial cues quickly, but with a natural delay that prevented a robotic feel. The total delay between the audio trigger and a noticeable reaction from the creature was 2.24 seconds, while the facial tracking had a delay of 2.36 seconds.

Research by Shiwa et al. [28] suggests that for users to perceive a system response as directly related to their action, the delay should ideally be less than two seconds. This aligns with the delays measured for both the audio and facial reactive behaviors, making sure visitors connect their actions with the creature's responses. In the specific case of the audio reactive response, the two second delayed response might seem too long in certain situations. This can be further worked on to make the delay smaller.

One expert from P&P Projects commented on the lifelike nature of the interactions: "The way it turns its head as if it's noticing you, and then follows you with its eyes… it really feels like it's aware of your presence. It's those small details that make a big difference."

7.2 Multi-layered behavior: (SHOULD HAVE)

A key requirement for the interaction engine is the ability to combine multiple layers of behavior, preventing a simplistic or repetitive character.

7.2.1 Observational analysis:

On the 14th of June 2024, at P&P projects the creature was assessed over a longer period. Special attention was put into observing the relation between the behavior layers. For this activity, ethical approval by the examination board was granted with a positive advice, with the application No. 240104. The information letter and consent forms can be found in the Appendix B.

These are the questions that were taken as a base to assess the multi-layered behavior:

Base Behavior Integration: Did the subtle movements of the base animation remain present and unobtrusive when layered with more complex behaviors? Did it contribute to a constant sense of life, even when the creature was engaged in other actions?

Default Behavior Variety: Did the randomization of the default behavior animations create a diverse range of movements and prevent predictability? Did it successfully convey a sense of autonomy and curiosity?

Behavioral Hierarchy: When both audio and facial cues were present, did the creature prioritize its responses appropriately, mimicking the shifting attention of a living being?

Moreover, movement graphs from the output of the creature were also taken as a base to assess the fluidity of the blending and motion. An example of one of such graphs can be seen in [Figure 30].

Base behavior (breathing)

Behavior composition (halve facial tracking halve sound

Blended Behavior + Base motion output towards the

Figure 30 Multi-layered approach.

7.2.2 Findings from this evaluation

The always-present base animation, simulating breathing and subtle movements, prevented the creature from ever appearing static, even when no other behaviors were active. The randomization of the default behavior animations proved successful in creating a diverse range of actions, preventing predictability, and conveying a sense of autonomy. One P&P expert remarked: "It's like it has its own little world going on".

The behavior logic successfully managed the hierarchy of responses. During the observation, a loud noise was introduced while the creature was tracking an expert's face. The creature immediately reacted to the sound, momentarily breaking its gaze, before smoothly transitioning back to following the expert. This fluid interplay between behavior layers reinforced the impression of a complex internal state.

The P&P Projects experts were particularly impressed with the multi-layered behavior, recognizing its contribution to a more captivating experience: "It's not just reacting to one thing... like some animatronics we've seen. It seems to be processing different stimuli and making choices about what to focus on. That's a real game-changer."

The observations confirmed that the multi-layered approach is effective in creating a lifelike animatronic character. The creature's constant base animation, the variety of its default behaviors, and its ability to prioritize and seamlessly blend between responses to different stimuli all contributed to a believable illusion of life.

7.3 Customization (SHOULD HAVE)

The ability to adapt the creature's behaviors to suit different theme park settings or character personalities was a crucial consideration.

7.3.1 Evaluating customization potential:

The following aspects of the interaction engine were examined to assess its capacity for customization:

The structure of the base behavior, default behavior, and reactive behaviors was designed to allow for easy replacement or addition of animation sequences. Key parameters within the behavior logic, such as reaction times, animation blending durations, and trigger thresholds, were designed to be adjustable. Moreover, the sensitivity of the facial tracking algorithm, determining how readily the creature responds to faces, can be fine-tuned to suit the environment and desired level of interaction.

This flexibility within the system's design makes it so that it can be adapted to meet the specific requirements of future projects. New animation sets can be created to reflect different character personalities, while adjustments to parameters like reaction times can dramatically alter the creature's perceived temperament.

7.3.2 Future considerations:

To fully use the customization potential, the development of detailed documentation will be prioritized in future iterations. This documentation would empower theme park designers and operators to tailor the creature's behavior or adapt it to other characters.

7.4 Reliability and robustness (SHOULD HAVE)

A big concern for any system deployed in a theme park environment is its reliability and ability to handle the demands of continuous operation.

7.4.1 Stability testing:

The interaction engine underwent testing over an extended period to evaluate its stability. This experiment took place at P&P projects from the 5th till the 7th of June 2024.

The system ran continuously for 48 hours in a controlled environment, with logs monitored for errors, memory leaks, or performance degradation.

The results of this extended test were positive. The system remained stable throughout, with no critical errors or unexpected shutdowns. After the time had elapsed, an evaluation of the logs was made. It was found that resource utilization remained consistent, so no memory leaks or inefficient code execution happened (which had been a problem in the past).

7.4.2 Error handling and recovery:

Equally important to stability is the engine's ability to handle unexpected events or errors gracefully so not to break the “magic”. To assess this, a series of fault injection tests were performed on the 7th of June 2024 at P&P projects. This test simulated scenarios such as loss of camera input, audio overload and communication errors.

The result from this test was that the system successfully detected the loss of input and automatically switched to a default behavior, preventing a complete shutdown or strange behaviors being sent to the character. Upon recovery of the camera feed, the interaction engine seamlessly resumed its normal operation. When subjected to excessively loud audio, the system did not stay in the audio loop indefinitely. Instead, the audio timer which detects a high level of noise for a big amount of time kicked in.

Simulating temporary disruptions in communication between the control computer and the servo controllers, made the servos not receive position commands. The result of this was that the character stopped moving. When re-establishing communication, the system automatically resynchronized. This means there was no need to reset anything. This is positive, but could still be improved by adding a emergency routine in the microcontroller so if the communication with the computer is lost, it can still play something so the character does not stay still.

7.5 Privacy, safety, and inclusivity (MUST HAVE)

Incorporating ethical considerations into the design and implementation of the interaction engine was a non-negotiable aspect of this project.

7.5.1 Privacy: minimizing data collection

The use of facial recognition technology, even without data storage, raises complex privacy concerns, especially within the strict regulatory environment of the EU. While the current system is designed to only detect the presence and location of faces, not for identification, for implementing this in a theme park further steps are needed to ensure compliance and ethical operation. A thorough Data Protection Impact Assessment (DPIA) would be crucial to identify and mitigate potential privacy risks. Additionally, the necessity and proportionality of using facial recognition technology must be carefully evaluated, ensuring that the benefits to guest experience outweigh the potential impact on privacy. Transparent communication with visitors, clearly outlining how the technology is used and obtaining explicit consent would be required for theme park deployment. In the Netherlands, facial recognition is prohibited unless the users have "given explicit consent for their data to be processed"[29].

7.5.2 Safety: prioritizing visitor well-being

Safety was also consistently assessed throughout the development process. The creature's range of motion was physically restricted to prevent any possibility of accidental contact with visitors.

Moreover, all animations, particularly those within the reactive behaviors, were designed to avoid any movements that could be perceived as aggressive or threatening. The focus remained on causing delight and wonder, never fear or discomfort.

7.5.3 Inclusivity: engaging a diverse audience

The interaction engine was designed to be as inclusive as possible, so that the creature could engage with a wide range of visitors. To accomplish this, the following was tested:

The chosen facial recognition library was tested with a diverse dataset of images to make sure it could detect faces across a spectrum of skin tones, ages, and genders. This was partially successful, as some of the faces were not detected. The specific issue seemed to lay in detecting faces with beards. In the future, a different facial recognition system can be chosen to make sure all faces are immediately recognized.

While not directly related to the interaction engine itself, the installation of the animatronic creature should be considered to make sure it is accessible to visitors of all abilities. This was acknowledged by the P&P experts and taken into consideration for future installation work.

7.6 User friendly interface (could have)

While a user-friendly interface for real-time customization was not achieved within the scope of this phase of development (categorized as a "could have"), an attempt to start building and experimenting with this interface was made, and can be seen in [Figure 32]:

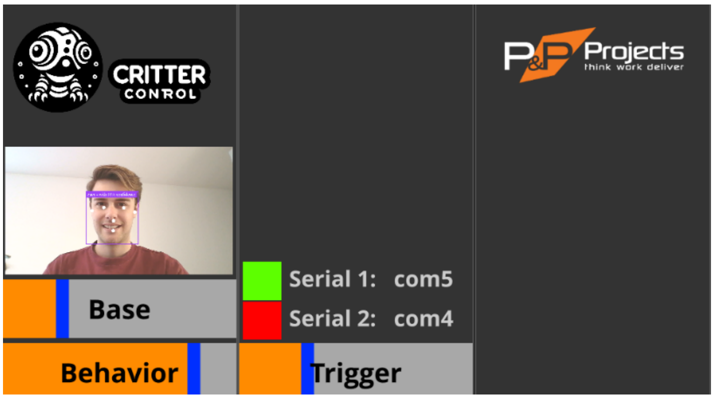


Figure 31: Interface attempt.

This interface includes the main sliders to blend the base and behavior composition, the camera feed to quickly adjust the field of view and an audio trigger sensitivity slider. These are important variables of the interaction engine that might require adjustments based on the setting and context the character is used on.

It also includes the connected ports for both microcontrollers connected to the PC. This is done to make sure that both controllers are receiving the position signals from the microcontroller. This interface is far from being completed, and it is a very rough draft as to what the final interface for the operator of this system could look like. More in-depth research could be made in the future to continue developing this interface.

7.7 Evaluation final findings

The evaluation results demonstrate that the developed multi-layered interaction engine successfully achieves the core objectives of creating natural, engaging, and ethically sound interactions for an animatronic character. With a combination of observation, quantitative metrics, and expert feedback, the system has proven its ability to create a believable illusion of life, responding dynamically to both human presence and environmental cues. The engine's modular design and adjustable parameters provide a strong foundation for future customization, allowing it to be adapted to diverse theme park environments and character personalities. The system's robust performance and error handling capabilities, combined with its privacy and safety measures, make a reliable and positive experience for both theme park operators and visitors alike. The resulting requirement table can be concluded as follows:

Table 3: Resulting requirement table.

Requirement	MoScoW	Accomplished
Natural interactions	Must have	Yes
Multi-layered behavior	Should have	Yes
Customization	Should have	Partially
User-Friendly interface	Could have	No
Reliability and robustness	Should have	Yes
Privacy	Must have	Yes
Safety	Must have	Yes
Inclusivity	Must have	Partially

8. Discussion and future work

8.1 Introduction

This project set out to design and implement a multi-layered interaction engine capable of filling an animatronic creature with a convincing illusion of life. The evaluation results show that this goal has been largely achieved. The engine's layered approach, incorporating subtle base animations, randomized default behaviours, and responsive reactions to audio and visual stimuli, creates a character that feels remarkably lifelike in its movements and interactions.

The blending between these behaviour layers contributes significantly to the overall believability of the creature. The smooth transitions allow actions to flow naturally, avoiding the jarring shifts that often plague simpler animatronic systems. This fluid behaviours, coupled with the creature's ability to prioritize responses based on the perceived importance of stimuli, creates the impression of a complex internal state, blurring the line between a mechanical puppet and a living being.

8.2 Customization and robustness

A key strength of the developed interaction engine lies in its modularity and potential for customization. While a user-friendly interface for real-time adjustments remains a goal for future development, the underlying architecture is inherently flexible. New animations can be easily integrated to reflect different character personalities, while parameters governing reaction times, blending durations, and trigger thresholds can be fine-tuned to tailor the creature's behaviour to specific needs. The Touch Designer platform makes it easy to make reconfigurations, add inputs or outputs and to customize the system to any other animatronic character.

The 48-hour stability test confirmed the system could handle continuous operation without performance degradation or critical errors, a crucial factor for deployment in a busy theme park. Moreover, the fault injection tests, simulating scenarios like the loss of camera input or communication errors, demonstrated the system's ability to handle unexpected events and recover seamlessly, preserving the illusion of life for visitors even in the face of technical hiccups. This is especially important in theme parks, which always strive to conserve the “magic" of their experiences.

8.3 Ethical considerations

The project aspired to create an engaging and ethically responsible experience, but the use of facial recognition, even without data storage, presents complex challenges. Navigating the privacy regulations of the Netherlands and the EU requires a more in-depth consideration of the technology's implications.

While the current system only detects the presence and approximate location of faces, a thorough Data Protection Impact Assessment (DPIA) is crucial to fully understand and mitigate potential risks. The DPIA should not only address technical aspects of data handling but also go into the ethical dimensions of using facial recognition in this context. For example, are there less privacy-invasive methods for achieving similar levels of interaction? Could the system be designed to function effectively without relying on facial recognition?

Moreover, for facial tracking transparency and user consent are mandatory in the EU [29]. Theme parks deploying this technology in the EU would need to clearly inform visitors about its use, outlining what data is processed and how and obtain explicit consent. This raises questions about how such consent would be obtained in a practical and user-friendly way within a theme park environment without breaking the “magic”.

Regarding inclusivity, even though during this project we could only identify potential discrimination with facial tracking towards individuals with a beard, OpenCV modules like the one used in this project tend to discriminate against user demographics [30]. This can be counteracted by using different approaches to detect visitors, such as other sensor technology or different facial tracking modules.

8.4 Impact on audience behaviour

This interaction engine could impact audience behaviour by transforming passive spectators into active participants. The P&P Projects experts, for example, were struck by the creature's ability to mimic the shifting attention of a living being, noting that "it's not just reacting to one thing...It seems to be processing different stimuli and making choices." This creates a powerful illusion of life, encouraging a different kind of interaction. Imagine kids whispering to a museum exhibit that listens and teaches them about the past, or families competing to get a smile from a character that responds to their movements. This interaction engine can create a more engaging and memorable experiences for visitors, encouraging them to spend more time interacting with the animatronic and creating a deeper connection with the storyline and environment being showcased.

Now this approach can be used for breathing life into static displays and characters. The multi-layered approach, combining subtle base animations with randomized default behaviours, offers a blueprint for creating a constant sense of life, even when a character isn't actively engaged in a performance. By incorporating triggers based on visitor proximity, sound, or even time of day, designers can create a dynamic experience that draws the audience in and makes them feel like they're witnessing a living, breathing being. The Touch Designer framework, with its visual scripting and possibility for customization, provides a relatively accessible platform for even those without extensive coding knowledge to experiment with these techniques.

8.5 Limitations and future work

Despite the project's successes, there are areas where further refinement and exploration are needed. One important limitation of the current system is the reliance on a rule-based approach for decision-making. While this provides a high degree of control and predictability, it comes at the cost of adaptability to unforeseen situations. The creature's responses are predetermined based on a finite set of rules, which may prove insufficient in handling the complexities and unpredictable nature of real-world interactions.

Another area for improvement lies in enhancing the facial recognition system's accuracy and inclusivity. While initial tests with a diverse dataset gave promising results, there were instances where certain faces were not reliably detected. This also shows the need for more refinement and optimization of the facial recognition, so that the creature can engage with visitors with all facial features equally.

Addressing the limitations of the current facial recognition system will be a priority in future iterations. This should involve testing with a diverse dataset of faces, including a big range of ages, ethnicities, and genders. Collaborating with experts in computer vision and machine learning will also be essential in optimizing the facial recognition algorithms, so that the system can accurately and reliably detect all faces. This commitment to inclusivity is crucial in creating a welcoming and engaging experience for every theme park visitor.

The successful development of this multi-layered interaction engine opens many opportunities for future research and development, with the potential to revolutionize the context of animatronic entertainment. Integrating artificial intelligence (AI) into the decision-making process holds immense promise for overcoming the limitations of the current rule-based system. This is done by the Disney's

Droids in Star Wars: Galaxy's Edge. Like the droids, by training AI algorithms on data collected from visitor interactions, the creature could learn to adapt its behaviour in real-time, responding in a more "Intelligent" way.

Moreover, an AI-powered system could analyse facial expressions, body language, and vocal cues to translate a visitor's emotional state and tailor its responses accordingly. This would allow the creature to engage in more meaningful interactions, going beyond simple reactions to showing a degree of understanding.

The current interaction engine relies primarily on audio and visual input. Expanding the creature's sensory capabilities could unlock new levels of engagement and realism. Integrating touch fabric such as Cang et all.[19] has explored, would allow the creature to respond to physical interaction, such as a gentle pat on the head or a handshake. This tactile feedback could trigger specific animations or serve as an input for the AI system as well.

As the current interaction engine relies primarily on audio and visual input, it limits the creature's ability to feel truly integrated into its environment. Imagine a scenario where the creature doesn't just react to a sudden loud noise but also shivers slightly in response to a cold breeze sensed by a temperature sensor, or playfully tracks a passing butterfly with its gaze using a more sensitive proximity detector. By expanding the creature's sensory inputs beyond sound and vision to include factors like temperature, wind, and even the presence of small, moving objects, a more immersive and believable experience can be achieved. This would create a stronger sense of shared environment between the animatronic and the audience, creating a feeling of being present in the moment together.

To fully realize the potential for customization, the further development of a user-friendly interface is crucial. This interface would empower theme park designers and operators to easily modify the creature's behaviours, adjust parameters, and create new animation sequences without requiring specialized technical knowledge. Moreover, an intuitive graphical interface to create new behaviours could allow operators to drag and drop animations, set trigger thresholds, adjust reaction times, and fine-tune the creature's overall personality. This would democratize the creative process, allowing for a wider range of individuals to contribute to the development of captivating and personalized animatronic experiences.

8.6 Closing

Ultimately, the true test of any animatronic system lies in its long-term performance within a real-world theme park environment. Deploying the creature in a controlled setting and gathering data on visitor interactions, system stability, and maintenance requirements will be essential. This long-term evaluation will provide valuable insights into the system's strengths and weaknesses, identifying areas for further optimization and informing the development of even more advanced and engaging animatronic characters in the future.

This page has been intentionally left blank

9. Conclusion

The ambition of this project was realized not just through lines of code and mechanical parts, but through the illusion of life within the animatronic creature. The true measure of success is not solely in the fluidity of its movements or the complexity of its responses, but in its capacity to make wonder and blur the lines between technology and nature. This project is a step in the landscape of animatronics, a move away from pre-programmed sequences towards a future with genuine interactions.

Another goal that was accomplished is the system's potential to go beyond its initial application. While conceived within the realm of theme park entertainment, imagine this technology breathing life into immersive theatrical performances, where animatronic actors share the stage with their human counterparts, responding to cues and improvising in real-time. Picture interactive museum exhibits where historical figures engage in lively dialogue with visitors, adapting their responses based on the flow of conversation. Envision these interactive companions offering comfort and stimulation to individuals facing social isolation, providing a bridge between technology and genuine human connection.

As artificial intelligence matures, its integration with animatronics holds the potential to revolutionize human-machine interaction. Imagine animatronic characters evolving beyond a fix set of rules and routines, capable of learning from each interaction and increasing the number of behaviours. Imagine a world where these creations surprise us with their emotional intelligence, responding not just to our words but also to our emotions and body language.

This vision, however, is not without its ethical obligations. As we go into this new era of sophisticated animatronics, we must proceed with mindfulness and a deep respect for the values that shape a just and equitable society. The act of mimicking life carries with it the responsibility to safeguard privacy, to ensure inclusivity, and to prioritize user safety above all else. Only then can this technology reach its full potential, not only as a spectacle, but as a force for positive change and meaningful human connection.

This project has demonstrated the immense potential of multi-layered interaction engines by answering the main research question. The combination of design, technical challenges, and ethical considerations has resulted in a system that not only captivates but also sets the stage for a new generation of interactive animatronics. Through continued research and development, we can unlock the full potential of this technology, creating experiences that blur the lines between fantasy and reality, sparking wonder and delight in audiences for years to come.

This page has been intentionally left blank

10. References

[1] C. A. Baker, “Creative choices and fan practices in the transformation of theme park space,” *Transformative Works and Cultures*, vol. 22, Sep. 2016, doi: 10.3983/TWC.2016.0974.

[2] B. Burns and B. Samanta, “Mechanical Design and Control Calibration for an Interactive Animatronic System,” in *ASME 2015 International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers Digital Collection, Mar. 2016. doi: 10.1115/IMECE2015-52477.

[3] B. Tomlinson, “Social characters for computer games,” *Interactive Technology and Smart Education*, vol. 2, no. 2, pp. 101–115, May 2005, doi: 10.1108/17415650580000036/FULL/XML.

[4] A. S. Krishen, Y. K. Dwivedi, N. Bindu, and K. S. Kumar, “A broad overview of interactive digital marketing: A bibliometric network analysis,” *J Bus Res*, vol. 131, pp. 183–195, Jul. 2021, doi: 10.1016/J.JBUSRES.2021.03.061.

[5] H. Knight and R. Simmons, “Layering Laban Effort Features on Robot Task Motions,” 2015, doi: 10.1145/2701973.2702054.

[6] D. Chi, M. Costa, L. Zhao, and N. Badler, “The EMOTE Model for Effort and Shape,” *SIGGRAPH 2000 - Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 173–182, Jul. 2000, doi: 10.1145/344779.352172.

[7] T. Ribeiro and A. Paiva, “The illusion of robotic life: principles and practices of animation for robots,” in *The seventh annual ACM/IEEE international conference on Human-Robot Interaction (HRI ’12). Association for Computing Machinery, New York, NY, USA*, 2012, pp. 383–390. doi: 10.1145/2157689.2157814.

[8] S. Kopp *et al.*, “Towards a common framework for multimodal generation: The behavior markup language,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4133 LNAI, pp. 205–217, 2006, doi: 10.1007/11821830_17.

[9] D. Reidsma and H. van Welbergen, “AsapRealizer in practice – A modular and extensible architecture for a BML Realizer,” *Entertain Comput*, vol. 4, no. 3, pp. 157–169, Aug. 2013, doi: 10.1016/J.ENTCOM.2013.05.001.

[10] H. van Welbergen, D. Reidsma, Z. M. Ruttkay, and J. Zwiers, “Elckerlyc A BML Realizer for continuous, multimodal interaction with a Virtual Human,” *Journal on Multimodal User Interfaces*, vol. 3, no. 4, pp. 271–284, Nov. 2009, doi: 10.1007/S12193-010-0051-3/METRICS.

[11] D. P. Davison *et al.*, “Things that Make Robots Go HMMM : Heterogeneous Multilevel Multimodal Mixing to Realise Fluent, Multiparty, Human-Robot Interaction,” 2017, doi: null.

[12] S. Wittig, M. Rätsch, and U. Kloos, “Parameterized Facial Animation for Socially Interactive Robots,” *Message Understanding Conference*, 2015, doi: 10.1515/9783110443929-051.

[13] J. A. Russell, M. Lewicka, and T. Niit, “A Cross-Cultural Study of a Circumplex Model of Affect,” *J Pers Soc Psychol*, vol. 57, no. 5, pp. 848–856, 1989, doi: 10.1037/0022-3514.57.5.848.

[14] G. Mehlmann, M. Häring, K. Janowski, T. Baur, P. Gebhard, and E. André, “Exploring a Model of Gaze for Grounding in Multimodal HRI,” 2014, doi: 10.1145/2663204.2663275.

[15] C. Mishra and G. Skantze, “Knowing Where to Look: A Planning-based Architecture to Automate the Gaze Behavior of Social Robots*,” 2022, doi: 10.1109/ro-man53752.2022.9900740.

[16] M. K. X. J. Pan *et al.*, “Realistic and Interactive Robot Gaze,” *IEEE/RJS International Conference on Intelligent RObots and Systems*, 2020, doi: 10.1109/iros45743.2020.9341297.

[17] W. S. M. Sanjaya *et al.*, "Design of Real Time Facial Tracking and Expression Recognition for Human-Robot Interaction," *J Phys Conf Ser*, 2018, doi: 10.1088/1742-6596/1090/1/012044.

[18] S.-H. Kim, H. Nam, S.-M. Choi, and Y.-H. Park, "Real-Time Sound Recognition System for Human Care Robot Considering Custom Sound Events," *IEEE* , 2024, doi: 10.1109/access.2024.3378096.

[19] X. L. Cang, "Towards an emotionally communicative robot : feature analysis for multimodal support of affective touch recognition," *T, University of British Columbia*, 2016, doi: 10.14288/1.0314100.

[20] A. Feng, Y. Huang, M. Kallmann, and A. Shapiro, "An Analysis of Motion Blending Techniques," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7660 LNCS, pp. 232–243, 2012, doi: 10.1007/978-3-642-34710-8_22.

[21] P. Joshi, W. C. Tien, M. Desbrun, and F. Pighin, "Learning controls for blend shape based realistic facial animation," *ACM SIGGRAPH 2003 Sketches and Applications, SIGGRAPH 2003*, Jul. 2003, doi: 10.1145/965400.965571.

[22] Y. Seol, J. Seo, P. H. Kim, J. P. Lewis, and J. Noh, "Artist friendly facial animation retargeting," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 6, pp. 1–10, Dec. 2011, doi: 10.1145/2070781.2024196.

[23] "Star Wars: Galaxy's Edge Droid Playtest." Accessed: Apr. 14, 2024. [Online]. Available: https://disneyconnect.com/star-wars-galaxys-edge-droids/

[24] "Disney introduces autonomous robots into Disneyland | blooloop." Accessed: Apr. 14, 2024. [Online]. Available: https://blooloop.com/theme-park/news/disney-vyloo-robots-disneyland/

[25] H. Knight and R. Simmons, "Layering Laban Effort Features on Robot Task Motions," 2015, doi: 10.1145/2701973.2702054.

[26] C. Mishra, C. Mishra, G. Skantze, and G. Skantze, "Knowing Where to Look: A Planning-based Architecture to Automate the Gaze Behavior of Social Robots*," 2022, doi: 10.1109/ro-man53752.2022.9900740.

[27] "torinmb/mediapipe-touchdesigner: GPU Accelerated MediaPipe Plugin for TouchDesigner." Accessed: Jun. 24, 2024. [Online]. Available: https://github.com/torinmb/mediapipe-touchdesigner

[28] T. Shiwa, T. Kanda, M. Imai, H. Ishiguro, and N. Hagita, "How Quickly Should a Communication Robot Respond? Delaying Strategies and Habituation Effects," *Int J Soc Robot*, vol. 1, no. 2, pp. 141–155, Feb. 2009, doi: 10.1007/S12369-009-0012-8/METRICS.

[29] "Home | Autoriteit Persoonsgegevens." Accessed: Jun. 24, 2024. [Online]. Available: https://autoriteitpersoonsgegevens.nl/en

[30] P. Terhorst *et al.*, "A Comprehensive Study on Face Recognition Biases Beyond Demographics," *IEEE Transactions on Technology and Society*, vol. 3, no. 1, pp. 16–30, Sep. 2021, doi: 10.1109/TTS.2021.3111823.

"During the preparation of this work, I used OpenAI Chat GPT, Grammarly, research Rabbit, Mendeley Referencing, Consensus and Google AI Studio for aid in formatting, grammar, spelling, brainstorming, proofreading and referencing. After using this tool/service, I thoroughly reviewed and edited the content as needed, taking full responsibility for the final outcome."

11. Appendixes

11.1 Appendix A – Microcontroller Code

```
#include <ESP32Servo.h>

// Define servo pins
int servoPins[] = {15, 25, 13, 12, 16, 18, 26, 33, 19, 32, 27, 17, 4, 21};
#define NUM_SERVOS (sizeof(servoPins) / sizeof(servoPins[0]))
Servo servos[NUM_SERVOS];

// Task handles
TaskHandle_t Task1;
TaskHandle_t Task2;

// Shared variables
volatile int servoIndex = -1;
volatile int servoPosition = -1;
portMUX_TYPE mux = portMUX_INITIALIZER_UNLOCKED;

// Function declarations
void TaskSerialRead(void *pvParameters);
void TaskServoControl(void *pvParameters);

void setup() {
  Serial.begin(115200);
  Serial.println("Setup started");

  // Attach each servo
  for (int i = 0; i < NUM_SERVOS; i++) {
    servos[i].setTimerWidth(16);
    servos[i].attach(servoPins[i]); // Specify pulse width for better servo control
  }

  // Create tasks on separate cores
  xTaskCreatePinnedToCore(
    TaskSerialRead,   /* Task function. */
    "Task1",          /* Name of the task */
    10000,            /* Stack size of task */
    NULL,             /* parameter of the task */
    1,                /* priority of the task */
    &Task1,           /* Task handle to keep track of created task */
    0);               /* pin task to core 0 */

  xTaskCreatePinnedToCore(
    TaskServoControl, /* Task function. */
    "Task2",          /* Name of the task */
    10000,            /* Stack size of task */
    NULL,             /* parameter of the task */
    1,                /* priority of the task */
    &Task2,           /* Task handle to keep track of created task */
    1);               /* pin task to core 1 */
  Serial.println("Setup complete");
}

void loop() {
  // Empty. Tasks are handling everything.
}

void TaskSerialRead(void *pvParameters) {
  char buffer[50];
  int len = 0;
  while (1) {
    if (Serial.available() > 0) {
      char incomingByte = Serial.read();
      if (incomingByte != '\n') {
        if (len < sizeof(buffer) - 1) {
          buffer[len++] = incomingByte;
        } else {
          len = 0; // Prevent buffer overflow
          Serial.println("Buffer overflow.");
        }
      } else {
        buffer[len] = '\0'; // Null-terminate the string
        char* token = strtok(buffer, ",");
        int index = atoi(token);
        token = strtok(NULL, ",");
        int position = atoi(token);
        if (index >= 0 && index < NUM_SERVOS) {
          portENTER_CRITICAL(&mux);
          servoIndex = index;
          servoPosition = position;
          portEXIT_CRITICAL(&mux);
        }
        len = 0; // Reset buffer length
      }
    }
  }
}

void TaskServoControl(void *pvParameters) {
  while (1) {
    if (servoIndex != -1) {
      portENTER_CRITICAL(&mux);
      int index = servoIndex;
      int position = servoPosition;
      servoIndex = -1; // Reset index after handling
      portEXIT_CRITICAL(&mux);
      if (index >= 0 && index < NUM_SERVOS) {
        servos[index].writeMicroseconds(position);
      }
    }
  }
}
```

11.2 Appendix B

11.2.1 Consent form

Consent Form for Breathing Life into Animatronics.

YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated [DD/MM/YYYY], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	○	○
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	○	○
Use of the information in the study		
I understand that the data collected from these interviews will be anonymized and used solely for academic research purposes, contributing to publications and presentations about the project.	○	○
I understand that personal information collected about me that can identify me, such as [e.g. my name or where I live], will not be shared beyond the study team.	○	○
I agree that my information can be quoted in research outputs	○	○
I agree that my real name can be used for quotes	○	○
I Consent to be audio recorded	○	○
Future use and reuse of the information by others		
I give permission for the expertise that I provide to be archived in an anonymised transcript so it can be used for future research and learning.	○	○

Signatures

_______________________ _______________________ _________

Name of participant [printed] | Signature | Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

__________________________ ____________________ _________

Researcher name [printed] Signature Date

Study contact details for further information:

(Removed)

Contact Information for Questions about Your Rights as a Research Participant

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee Information & Computer Science: **(Removed)**

11.2.2 Information letter

To: [Name of Expert]

From: Jorge Davo Sainz

Date: [Date]

Dear [Name of Expert],

Thank you again for agreeing to share your valuable expertise for my graduation project, "Breathing life into animatronics" I am incredibly grateful for your willingness to contribute to this research.

As we approach our scheduled interview on [Date] at [Time] via [Location/Platform], this letter aims to provide some background information about the project and the areas we'll be discussing.

My project centers on developing a multi-layered interaction engine for an advanced animatronic creature, aiming to create an immersive and engaging experience for theme park visitors. This engine will use various sensor inputs like facial recognition, audio detection, and proximity sensors to drive the creature's real-time responses, fostering a sense of life and personality. Think of the engine as the main system that runs the animatronic character, its brain.

Our conversation will primarily revolve around your expertise the theme park industry. This interview will be semi-structured and will be audio recorded. Written notes will also be taken during this interview. The data collected from these interviews will be anonymized and used solely for academic research purposes, contributing to publications and presentations about the project.

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee Information & Computer Science: **(Removed)**

Your insights will be vital for shaping the project's direction and collaborating the creation of a truly innovative and engaging animatronic experience.

I am eager to learn from your experience and discuss your perspectives on this project. Please don't hesitate to reach out if you have any questions beforehand.

Thank you once again for your valuable contribution.

Sincerely,

Jorge Davo Sainz

Creative Technology

University of Twente

11.3 Appendix C

Creature animatronic animation state list

Default states:

	Animation	Story	sound
01	Breathing	Random breathing	Breathing sound
02	Looking around	Examining its surroundings	Mumbling
03	Shaking head	Like a dogs stress relief	
04	Sneezing	Itching nose	Sneeze
05	Ticking object on ground	Tries to open it?	Irritated growling
06	Sighing	It is getting bored	Deep exhale
07	Pride	Looking/studying object with tilted head	Mumbling interested
08	Reaction on object	Object is glow up	Object sound? Excitement
09	Head tilt + ear movement	Ear itches	
10	Power nap	Falls asleep and quick wakeup	Snorring
11	Hick-ups	Has the hick-ups	Sweet high pitched 'Hick'
12	Tired	Yawning and stretching, smacking mouth	Deep inhale and exhale
13	Excitement	Gasping eyes wide open	Gasping
14	Burping	Uncomfortable little burp	Burp!
15	Sniffing	Sniffing the object he is holding	Sniffing sound
16	Sharing	Wants to show his glowing object to someone	

Detection / Interrupt states:

	Animation	Story	Sound
01	Scared shaking	Startles to sudden close frontal movement	Screams
02	Interested follow	Detected person wakling by	Humming
03	Giggling	When tickeling its paws	Giggeling
04	Blinking fast	(phone) camera flashlight	Irritated growling
05	Follow peoples hand	Hand motion follow	Interested mumbling
06	Gets angry	Stealing 3rd object from set (RFID)	Angry Growling
07	Parroting	Recognizes sounds/voices	Repeat sound
08	Follow object	Follows close frontal objects	Interested mumbling
09	Begging	Food recognition / use Acorns on set foot (RFID)	Hmmmmmm
10	Sadness	After interaction with person ends	Sighing sound
11	Disgust	An object is held close -> sniffs-> disgust	Bleeeeerk!
12	Shyness	When a person is getting too close to creature	
13	Gratitude	Audio recognition on Bye!, Goodbye,	Saying bye in own dialect
14	Laughing	Audio recognition -> reaction on people laughing	Laughing