CONVERSATIONAL AGENT EXPRESSING OCEAN PERSONALITY AND EMOTIONS USING LABAN MOVEMENT ANALYSIS AND NONVERBAL COMMUNICATION CUES

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

COMPUTER ENGINEERING

By

Sinan Sonlu

August 2019

CONVERSATIONAL AGENT EXPRESSING OCEAN PERSONALITY AND EMOTIONS USING LABAN MOVEMENT ANALYSIS AND NONVERBAL COMMUNICATION CUES

By Sinan Sonlu

August 2019

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Uğur Güdükbay(Advisor)

Hamdi Dibeklioğlu

Yusuf Sahillioğlu

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

CONVERSATIONAL AGENT EXPRESSING OCEAN PERSONALITY AND EMOTIONS USING LABAN MOVEMENT ANALYSIS AND NONVERBAL COMMUNICATION CUES

Sinan Sonlu
M.S. in Computer Engineering
Advisor: Uğur Güdükbay
August 2019

Conversational human characters are heavily used in computer animation to convey various messages. Appearance, movement and voice of such characters influence their perceived personality. Analyzing different channels of human communication, including body language, facial expression and vocalics, it is possible to design animation that exhibit consistent personality. This would enhance the message and improve realism of the virtual character.

Using OCEAN personality model, we design internal agent parameters that are mapped into movement and sound modifiers, which in turn produce the final animation. Laban Movement Analysis and Nonverbal Communication Cues are used for the operations that output bone rotations and facial shape key values at each frame. Correlations between personality and spoken text, and relations between personality and vocal features are integrated to introduce compherensive agent behavior.

Multiple animation modification algorithms and a personality based dialogue selection method is introduced. Resulting conversational agent is tested in different scenarios, including passport check and fastfood order. Using a speech to text API user controls the dialog flow. Recorded interactions are evaluated using Amazon Mechanical Turk. Multiple statements about agent personality are rated by the crowd. In each experiment, one personality parameter is set to an extreme while others remain neutral, expecting an effect on perception.

Keywords: conversational agent behaviour, emotion, computer animation, personality, laban movement analysis, nonverbal communication.

ÖZET

LABAN HAREKET ANALİZİ VE SÖZSÜZ İLETİŞİM İŞARETLERİ KULLANARAK OCEAN KİŞİLİĞİ VE DUYGU SERGİLEYEN KONUŞABİLEN ARACI

Sinan Sonlu
Bilgisayar Mühendisliği, Yüksek Lisans
Tez Danışmanı: Uğur Güdükbay
Ağustos 2019

Konuşabilen insan karakterler bilgisayar animasyonunda çeşitli mesajları iletme amacıyla sıklıkla kullanılmaktadır. Bu karakterlerin görünüşü, hareketi ve sesi kişiliklerinin algılanmasında etkili olmaktadır. Beden dili, yüz ifadesi ve seslendirme gibi insan iletişiminin farklı kanallarını analiz ederek, tutarlı bir kişilik sergileyen animasyon tasarlamak mümkündür. Bu işlem mesajı pekiştirip sanal karakterin gerçekçiliğini artıracaktır.

OCEAN kişilik modelini kullanarak, sırasıyla son animasyonu oluşturan, hareket ve ses dönüştürücülerine eşlenen dahili aracı parametreleri tasarlıyoruz. Laban Hareket Analizi ve Sözsüz İletişim İşaretleri, her karede kemik dönme açıları ve yüz şekil anahtar değerlerini üreten işlemler için kullanılır. Kişilik ve konuşulan metin arasındaki ilişkiler ve kişilik ile ses özellikleri arasındaki bağlantılar bütüncül aracı davranışı için ilave edilmiştir.

Birden fazla animasyon dönüştürme algoritması ve kişilik tabanlı bir diyalog seçme metodu tanıtıldı. Geliştirilen konuşabilen aracı pasaport kontrolü ve hazır yemek siparişi dahil farklı senaryolarda test edildi. Kullanıcı sesi yazıya dönüştüren bir uygulama programlama arayüzü sayesinde diyalog akışını kontrol eder. Kayıt edilen etkileşimler Amazon Mechanical Turk sistemi kullanılarak değerlendirildi. Aracı kişiliğini hakkında çoklu ifadeler toplum tarafından derecelendirildi. Her deneyde, bir kişilik parametresi uç noktaya ayarlanırken diğerleri nötr kalır ve algı üzerinde bir etki beklenir.

Anahtar sözcükler: konuşabilen aracı davranışı, duygu, bilgisayar animasyonu, kişilik, laban hareket analizi, sözsüz iletişim.

Acknowledgement

I would like to thank Prof. Dr. Uğur Güdükbay for his guidance throughout my research.

I would also like to thank to the members of the jury, Asst. Prof. Dr. Hamdi Dibeklioğlu and Assoc. Prof. Dr. Yusuf Sahillioğlu, for their insightful comments and valuable questions.

I also would like to thank my family and friends for their support.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

1.1 Context and Motivation

Computer animation is an ever-growing field used for many different tasks in entertainment and education, including animated movies, video games, and interactive assistants. Virtual human characters are heavily used in computer animation to communicate with the viewer. By analyzing the elements of an animation that transfer the message to the viewer, communication could be enhanced. Multiple elements such as the environment and virtual characters in an animated scene contribute to the message. Each element could express some information in various dimensions. For example, a slanted posture in a character animation could express sadness (emotion), tiredness (physical state) or introversion (personality), and may signal a desk job (occupation) or scoliosis (health condition). Each scene element requires an analytical approach to comply with the intended message, yet the most important one is the human character.

Humans perform regular interpersonal communication through multiple channels of facial expression, body language, voice, verbal style, and verbal content [1]. The average viewer expects to see the same transmission in virtual characters or agents if artificial intelligence is in charge of the behavior. The lack of transmission in certain channels would render the agent lifeless and unnatural, disturbing the message. Furthermore, the lack of coordination between different channels could cause confusion and contradiction.

It is possible to modify elements of an animated scene to enhance the message. Adding personality to the agent would improve its effectiveness and supply a basis to manage related animation modifications. Previous research indicates the correlation between facial features and personality inferences [2]. Physical appearance and body posture also influence people's judgment of other's personality [3], as well as speech [4] and dialogue content [5]. Similarly, the correlation between agent gestures, speech content, and personality in animation is recorded [6]. Furthermore, we can modify movements to alter agent personality [7].

It is possible to record all channels of human communication during an act using expensive motion capture equipment and specially trained actors, however, the resulting animation is limited to one mood and personality. Even with a modular approach, the recording and storing animations for every possible combination is infeasible.

With a procedural approach, we can analyze an existing animation for different channels of transmission to control the message. We can modify an existing transmission in a channel to express the desired personality and emotion and generate a non-existing transmission. The closer to the regular human communication, the more realistic and natural the animation would look and sound.

With these observations, we introduce a conversational agent system that expresses personality and emotions through multiple channels of body language, facial expression, verbal style, and voice. The agent is capable of giving appropriate responses to user input based on predefined scenarios.

1.2 Introducing the Framework

In technical terms, we translate agent personality features into automated modifiers for three-dimensional (3D) animation, speech content, and auditory parameters. We borrow theories from psychology and human sciences to analyze and parametrize each concept.

We use the OCEAN model [8] to describe the personality of an agent. Although this theory is built for humans, it is also applicable to virtual personalities that represent real humans. There exist several extensions for personality in animals [9, 10]. These could yield interesting results related to nonverbal behavior in humans and could be useful for implementing non-human agents.

Although various personality factors in the OCEAN model include explanations that could be interpreted into movement, we use Laban Movement Analysis (LMA) [11, 12] to better parametrize the motion. LMA is introduced by Rudolf Laban has its roots in dance choreography. However, it is generalized to all human motion and used by many researchers, including [7, 13, 14, 15, 16]. LMA includes components to analyze motion in terms of spatial and temporal relations of body parts. These concepts could be translated into computer language using computable descriptors for human motion [17].

As a link between the OCEAN model and LMA, we make use of various sources on nonverbal communication [18, 19, 20] and personality [21], as well as the previous research that connects these concepts together [22, 7, 23, 24].

To include personality to the agent's speech, we make use of a theory introduced by Mairesse et al. [5], [25] that connects speech content and OCEAN personality. They introduce correlations between a set of linguistic cues [26] and various personality types. We adopt the introduced theory to handcraft dialogue text that fits each personality type. To adjust the vocal features of speech that influence the perceived personality, we use the mapping introduced by Polzehl [4].

1.3 Contributions

- We develop a method for analyzing and modifying Laban Shape Quality of an animation, based on relative distances of target limbs to multiple anchor points through time.
- We develop a technique to choose alternative dialogue text based on the agent personality and a method to craft OCEAN text alternatives.
- We propose a mapping between facial expressions and OCEAN personality types based on the results of the related experiment.
- We introduce a conversational agent architecture that uses human communication channels to express different messages based on predefined scenarios. The system is capable of expressing personality through animation modification, dialogue selection, and voice generation, as well as expressing emotions through facial expressions based on selected dialogue and personality.

1.4 Organization of the Thesis

Following the introduction, Chapter 2 summarizes the necessary background, including communication, personality, and movement models, in addition to mentioning related works. Chapter 3 explains the implemented architecture and describes the flow of the modification process. Chapter 4 elaborates on each movement modification process, describing the algorithms and techniques used in Laban Shape, Laban Effort and Nonverbal Communication modifications. Chapter 5 explains how personality is mapped into dialogue selection and movement modification parameters, as well as emotion mapping. Chapter 6 introduces the scenarios that are used in the experiments. Chapter 7 contains the evaluation and results of experiments performed with a user study. Chapter 8 concludes the thesis with a reflection on the results and a list of possible future research topics.

Chapter 2

Background and Related Works

2.1 Communication Model

Communication is the task of a message exchange between a sender and a receiver. An immediate transfer of the intended message is not possible, therefore it is encoded at the sender and transferred through multiple channels. The receiver requires a decoding process to understand the message. According to [1], human communication channels are grouped as facial expression, body language, voice, verbal style, and verbal content, as seen in Figure 2.1.

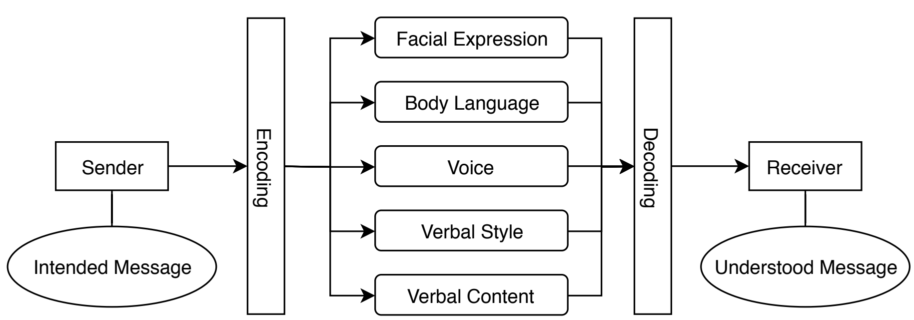


Figure 2.1: The communication channels in human interaction.

A more comprehensive model by [27] groups elements of interpersonal communication into two main categories of direct and indirect communication. Indirect communication includes all external factors to sender and receiver. This corresponds to the surrounding environment. For example, a dinner setup would have different influences on the message compared to a business meeting. Passive objects, lighting, ambient sounds, and music are examined under this category.

Direct communication is internal to the sender and the receiver, and it includes anything that belongs to them or emerges from them. Direct communication is categorized as *verbal* and *nonverbal* [27] (see Figure 2.2). The verbal category covers language-related topics, including grammar and word preferences, that form the message as a sentence. This sentence could be spoken with sound, or transferred using a language substitute such as American Sign Language.

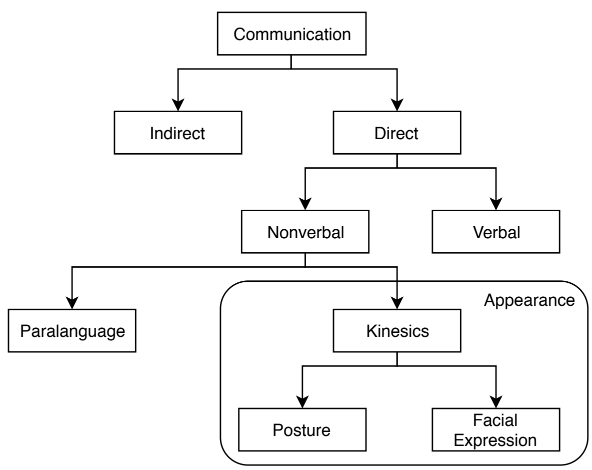


Figure 2.2: The categories of communication.

The nonverbal category covers *paralanguage* and *kinesics* [27]. Paralanguage is defined as the articulation of the vocal apparatus. It covers non-speech sounds, such as "umm" and "shhh", vocal features, such as tempo, pitch, and quality, and tonation, i.e., stressing some parts of speech for indication). Kinesics include all movement resulting from the muscular and skeletal shift. It is encapsulated by appearance, which includes body shape, clothes, hairstyle, and accessories. The muscular activity in the face is examined in a separate group called facial expression, and the remaining movement is examined in posture.

In terms of its function, nonverbal communication elements are examined as lexical, descriptive, reinforcing, embellishing and incidental, as seen in Figure 2.3.

Lexical elements have meaning by themselves without the aid of speech, like a hand gesture that means “come here” or making a “shhh” noise to mean “be quiet”. Descriptive elements are illustrative. For example, drawing a shape in the air to explain size or shape, or pointing to indicate position, or mimicking a noise is descriptive.

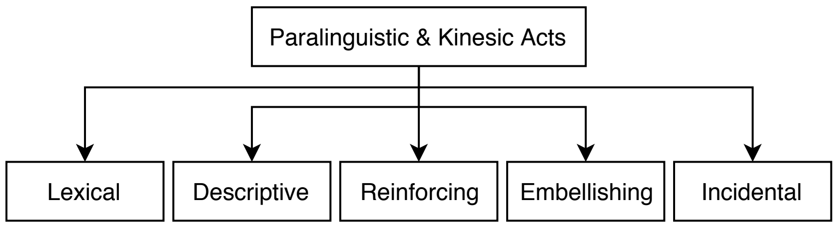


Figure 2.3: The types of paralinguistic and kinesic acts.

Reinforcing elements emphasize or highlight the verbal group, they augment the speech to make it more impactful, like clenching the fist to accompany an angry statement. Embellishing elements also accompany the speech, but without a specific goal to reinforce or describe. Key [27] notes that embellishing elements coordinate with personality types and speech would be dull and unattractive without these elements. Incidental elements do not contribute to the current dialogue. They are coexisting elements, like walking to a place while talking to a friend.

2.2 Communication in Computer Animation

Animation is the illusion of movement through temporal sequential images. Computer animation includes interpolated positions and rotations that are controlled by keyframes at different time points. An object-oriented approach to animation is possible with computers. Computers could also use automated methods to generate animation.

We use the term “scene” to mean a unit of animation that expresses a message to the viewer. The duration of a scene could be as long as the message requires. Since we focus on conversational agents, we assume at least one agent is present in any scene. Based on the communication model summarized in the previous section, we introduce the animated scene model shown in Figure 2.4. Based on this model, a scene consists of an environment and at least one agent. The environment covers scene properties, lighting, camera, and sound. It corresponds to indirect communication elements in [27]. The properties include all passive objects in the scene. Lighting influences scene appearance, the camera determines visibility and proximity, and sound includes music and ambient effects. In our implementation, the environment changes per scenario, and we do not use it to express personality.

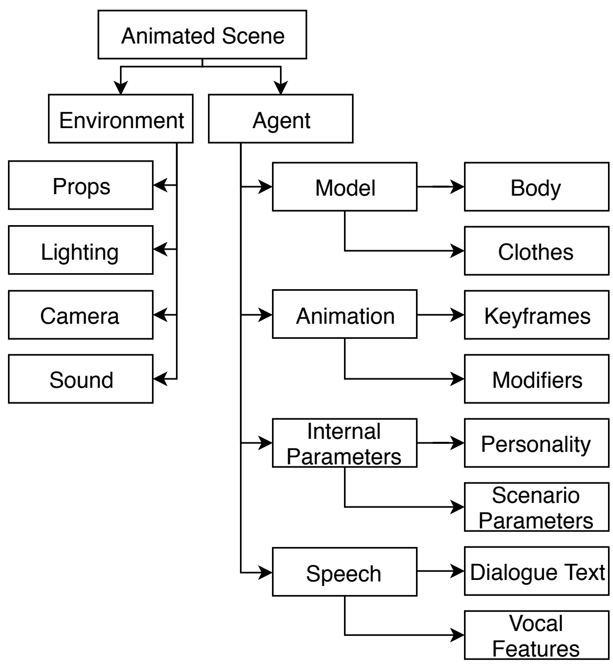


Figure 2.4: The components of an animated scene.

The agent appears on screen through its model. Body and clothes are examined under the model category, and they both passively influence the perceived personality. Throughout multiple scenes, the body stays the same, but clothes can change to influence the message. We use different models to find a neutral one in terms of its personality, and change its clothes based on the scenario.

The animation of the agent is stored as bone rotation keyframes. An agent could possess many different animation clips, yet only one could play at a given time. The agent has animation modifiers that alter the base rotations at each frame to produce a pose that expresses some personality. We examine both keyframes and modifiers under the animation category, which corresponds to kinesics in the communication model. Active keyframes, or in other words base animation, could be of type lexical, descriptive or reinforcing; on the other hand, the function of modifiers is to add embellishing elements to express personality. Internal agent logic chooses the base animation to accompany its dialogue, and animation modifiers use multiple parameters determined by agent personality.

Internal parameters of the agent include personality as well as scenario-specific information such as name, surname and passport data. Speech is examined under agent, and it contains dialogue text that corresponds to verbal elements and vocal features that corresponds to paralanguage. All these components contribute to the message of an animated scene, and the perception of the message is strongly related to the agent personality.

2.3 OCEAN Personality Model

Five-factor Personality Model, also known as Big Five Personality Traits, is a commonly used personality classification framework in psychology. In this model, the personality of an individual is analyzed in five orthogonal dimensions including Openness to Experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism, that form the acronym: OCEAN.

Each dimension is two ended, and the positive and negative ends are explained with multiple personality attributes that can be grouped in one factor. Although originally it is a problem of classification, meaning a person is predicted as an introvert or an extravert without a magnitude for it, using a continuous value for each factor has convenient implications for calculating movement parameters. Because of this, there are multiple examples in computer science that use continuous OCEAN parameters. This also makes it possible to compare two individuals with the same factor classification.

The following subsections summarize each OCEAN factor, including a short description from the psychology literature [21, 28], related nonverbal communication cues [19, 18], related speech features [5], and related vocal features [4]. Although there is not a comprehensive theory that links OCEAN personality features into nonverbal behavior, traits that personality factors encapsulate are directly related to nonverbal communication cues. For example, a shy person would look down to avoid eye contact, and being shy is a trait of negative extraversion, therefore we associate looking down with negative extraversion. We use resources on body language and nonverbal communication to find such links.

2.3.1 Openness to Experience

Openness indicates the intellectual aspect of the character. On the positive direction, the individual is associated with curiosity, imagination, adventurousness, attentiveness to emotions, preference for variety and awareness for aesthetics. People with negative openness tend to be conventional. They prefer familiar routines and do not have a wide range of interests. Such individuals often neglect emotions and aesthetics.

Positive openness is reflected in a welcoming body language. Hands tend to be open, and palms facing up and towards the listener. Shoulders tend to be relaxed. Feet have a larger angle in between. On the other hand, negative openness uses a protective body language. Hands tend to be closed, showing the backside to the listener. Shoulders are tense and shifted towards up. Feet are closer.

According to Mairesse, people with positive openness use many positive emotion words, many politeness forms, high meaning elaboration, and complex constructions, and they prefer less frequent but longer words in speech. Mairesse's speech parameters are the opposite of the negative side for each OCEAN factor.

The vocal features of positive openness are associated with high relative pitch, high pitch variation, varying tempo, and sonorous voice quality. Negative openness is associated with low relative pitch, low pitch variation, high tempo, and tense voice quality, resulting in a monotonous impression.

2.3.2 Conscientiousness

Conscientiousness indicates the carefulness and diligence of the individual. Positive conscientiousness is associated with efficiency, being organized, self-disciplined and dependable. Conscientious people tend to aim for the achievement and act in a planned manner. On the other side, individuals with negative conscientiousness are easy-going and disordered. They tend to be careless, forgetful and unreliable.

Body language of a high conscientiousness individual tends to be controlled and not in a hurry. A straight posture is kept, with shoulders shifted back. The unintentional movement of the limbs is very rare. On the opposite side, negative conscientiousness emerges as a hurried clumsy movement in the body. The posture is slanted, shoulders are shifted front and limbs tend to fluctuate. Attention span is low, which could reflect as the frequent head turns.

Mairesse finds positive conscientiousness individuals use few and straight to the point words with a formal language. They have few disfluencies and use fewer references to friends. They tend to check whether the information is conveyed correctly. Negative conscientiousness emerges as vague, careless, many words with frequent negations and disfluencies. They tend to use many filler words.

High conscientiousness individuals speak with a slower tempo with many pauses. Their voice quality is tense, relative pitch is medium and overall impression is calm. Low conscientiousness individuals' vocal features include higher tempo, relaxed voice quality, lower relative pitch.

2.3.3 Extraversion

Extraversion indicates the social aspect of the individual. Many personality theories include a variation of this factor; therefore, it is one of the most analyzed and easy to differentiate dimensions. People with high extraversion tend to be outgoing, energetic, talkative, warm, spontaneous and optimistic. They tend to focus more on the outside world, therefore enjoy being around people. They act instead of thinking deeply and are generally happier. People with negative extraversion (called introvert in the literature) tend to be shy, quiet, passive, moody. They are more focused on the mental self, neglecting physical outside elements. They enjoy solitary activities more and do not enjoy being around other people, therefore they appear unhappy and less energetic.

Body language of positive extraversion is spreading and relaxed. Extraverts tend to place their feet with wider space in between. They move a lot, and they tend to keep their hands farther away from the torso. They keep their heads raised and make eye contact most of the time. Negative extraversion is associated with compressing posture, such individuals tend to keep their hands closer to the torso. Their limb movement is lesser, and they keep their heads low avoiding eye contact. Shoulders are shifted up slightly, suggesting tension.

In a speech, according to Mairesse, positive extraversion is associated with pleasure talk and compliments. They use positive emotion words, and they tend to think out loud. They use simple constructions with few pauses, and their language is informal. Extraverts tend to use more swear words, and they exaggerate topics. They use shorter words, and their vocabulary is poor. On the opposite side, negative extraversion is associated with problem talk and elaborated constructions. Such individuals tend to use a formal language with many negations and negative emotion words. They have many pauses and use longer words. They have a rich vocabulary and they use realistic descriptions.

Vocal features of positive extraversion are high relative pitch, high pitch variation, high tempo and relaxed voice quality resulting in a lively overall impression. People with negative extraversion tend to speak with medium relative pitch and medium pitch variation. Their voice quality is damped, which makes the overall impression calmer.

2.3.4 Agreeableness

Agreeableness indicates the tendency for cooperation and being in harmony with others. People with positive agreeableness appear kind, sympathetic, trustworthy, helpful and considerate. They appear more friendly to others, while negative agreeableness is associated with selfishness and lack of empathy. Individuals with low agreeableness tend to be malicious, suspicious and uncooperative. They could manipulate others to reach their goals and they have little interest in other's problems.

People with positive agreeableness tend to make slower movements. They usually tilt their heads sideways, showing interest. Their movement is lighter and more flexible, adjusting to the environment in a better way. People with negative agreeableness tend to make sudden movements. They avoid eye contact and hide their hands most of the time, resulting in a closed body language.

Mairesse's research relates positive agreeableness with using more empathy words and few negations and swear words. Such individuals use longer but frequent words and use many acknowledgments. Their speech includes agreement and positive feelings most of the time. People with negative agreeableness use many negations and lack politeness forms. They use shorter words and have many pauses. Their talk is focused on problems and dissatisfaction.

Vocally, positive agreeableness uses low relative pitch with medium pitch variation and tempo that emphasizes calmness. Negative agreeableness is differentiated with its high tempo and tense voice quality.

2.3.5 Neuroticism

Neuroticism indicates the tendency of having emotional breakdowns. While other factors are associated with positive values in the positive direction, neuroticism is reversed. People with positive neuroticism tend to be worried and easy to upset. They experience dramatic mood changes and often feel anxiety, depression, and loneliness. They are more self-conscious and they appear more vulnerable. Individuals with negative neuroticism are emotionally stable, calm and confident. They rarely feel sad or depressed, and they are generally relaxed.

Individuals with positive neuroticism display an unpredictable body language. They could use out of place gestures and expressions in the extreme case. Faster limb fluctuations are visible in such individuals. They move quickly and shift attention frequently. Shoulders are lifted and legs are close together, resulting in a tense posture. Negative neuroticism is associated with proper body language, such individuals express slower and more precise movement. Their attention span is longer and body posture is relaxed.

According to Mairesse, positive neuroticism is related to polarized content, frustration, and disfluencies in speech. Such individuals use many self-references and focus on problems. Their claims are direct and they use frequent words with many filled pauses, which are spoken words or sounds that are used for filling gaps during a talk, such as "I mean", "err", "um", and "well". Negative neuroticism is associated with neutral content and pleasure talking. Such individuals have few self-references and fewer disfluencies. They show less frustration and their descriptions are more realistic.

Neuroticism does not introduce much difference in vocal features. People with high neuroticism tend to have lower relative pitch and fragile voice quality, whereas negative neuroticism is associated with calmness and sonorous voice quality.

2.4 Laban Movement Analysis System

Laban Movement Analysis (LMA) is a framework used for describing human movement in different dimensions [11, 12]. The main categories of LMA include *Body*, *Effort*, *Shape*, and *Space*, as seen in Figure 2.5. Body examines the interrelationships of limbs. Effort examines subtle characteristics of movement concerning inner intention. Shape examines the change of body during movement. Space examines the motion about the environment. We use Effort and Shape components to analyze and modify the animation. The space component is not used because we assume the agent does not change its position during a conversation. The body component corresponds to animation keyframes and it is not useful during modification.

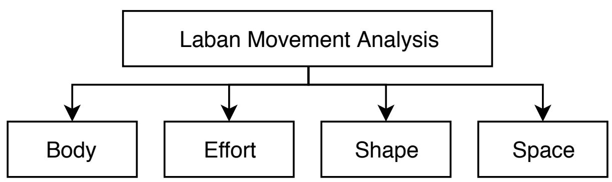


Figure 2.5: The categories of Laban Movement Analysis.

The effort component analyzes the movement in four different aspects of *Weight*, *Space* (not to be confused with the category Space), *Time*, and *Flow*, as shown in Figure 2.6. Each effort has two opposite polarities.

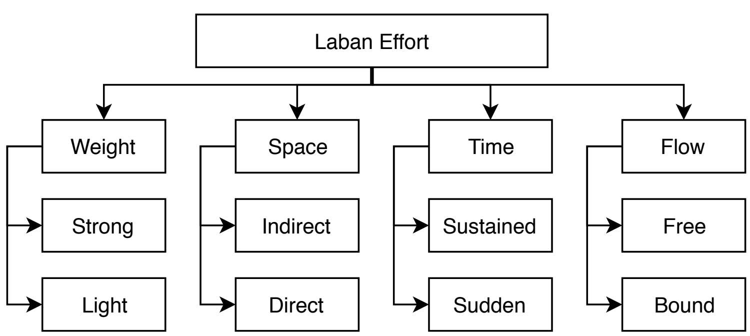


Figure 2.6: The aspects of Laban Effort component.

Weight Effort describes the attitude of using the impact of body weight for a purpose. It could be light or strong. Light weight is associated with delicate, gentle and soft movements, whereas strong weight is associated with forceful, powerful and demanding ones. Space Effort describes the consciousness towards surrounding space, and it could be indirect or direct. Indirect space is related to scanning, flexible motion; and direct space is related to linear, single-focused motion.

Time Effort describes the attitude towards the passage of time. It could be sustained or sudden. Sustained time is associated with reluctant, prolonging, lingering motion; on the other hand, sudden time is associated with quick, urgent, rushed motion. Flow Effort describes the continuous quality of motion. It can be free or bound. Free flow is related to loosening, fluent and careless movement; while the bound flow is related to restrained, tight and careful movement.

Shape category contains four subcategories: *Shape Forms* describe how the bodies of static shape could look like. The modes of *Shape Change* describe the interaction of the body with itself and the environment. *Shape Qualities* describe the change of the body towards points in space. *Shape Flow Support* describes the change of torso to support the movement of limbs. We use *Shape Qualities* in the implementation to analyze the attraction of limbs toward points in all six directions around the body. Based on the attraction, the movement could be rising or sinking, spreading or enclosing, advancing or retreating.

2.5 Similar Works

Similar works could be grouped into two different categories, based on how they aim to express personality or emotion: Using movement and using text or speech. Although we aim to output an act that expresses personality and emotions, there are works worth mentioning that receive movement data as input to predict personality. Such research would help to explore the link between movement and perceived personality, benefiting both analyzation and generation. Ran et al. [29] introduce an automated model to predict Laban qualities of recorded motion, using Microsoft Kinect and Certified Movement Analysts to label the dataset. They use features such as relative limb distances, the center of mass positioning, limb speed and acceleration, and directedness of distal limbs for Multi-task Learning with Elastic Net regularization. Camurri et al. [30] capture a dancer performing the same dance by expressing four different emotions, and surveys observers to introduce an automatic recognition model of perceived emotion. Aristidou et al. [31] propose a framework to extract Laban motion qualities, and automatically predict the included emotion as one of the 12 emotion states using Random Forests, Extremely Randomized Trees and Support Vector Machines.

2.5.1 Expressing Personality and Emotions Using Movement

These works generate or modify the movement of an articulated figure to express personality. Smith and Neff [32] focus on gesture performance using various motion adjustment parameters mapped into OCEAN personality. They adopt thin-slice psychology that indicates people can make personality judgments based on limited data in a quick manner with high accuracy. With a user study, they record each motion parameter's effect on perceived personality in short animated clips. They conclude people's perception of personality happens in a two-dimensional (2D) space when it is only based on the gesture. They introduce these dimensions as plasticity (openness and extraversion) and stability (conscientiousness, agreeableness, and neuroticism). With a second study, they verify the link between particular motion parameters and plasticity-stability perception. They include the speech of the motion capture actor in one experiment; however, they do not investigate the influence of spoken text and speech style. They use realistic human models with hidden faces and a wooden figure with no face in the experiments.

Durupınar et al. [7] make use of links between the low-level motion parameters and Laban Movement Analysis to express OCEAN personality in human figures performing atomic actions. They focus on the whole-body motion and introduce keyframe manipulation to compensate time-based Laban Movement parameters. Instead of defining motion parameters per Laban Effort factor, they implement Laban Drives, which are combinations of various Laban Effort elements, and collaborate with Laban Movement Analysis experts for adjustments. In the user study, they compare the OCEAN personality of two figures that have different Laban Drives and ask which figure expresses a particular OCEAN factor more. They achieve consistent results for the perceived personality of male, female and wooden figures with different base animations. As a result, they introduce a mapping between OCEAN and Laban parameters.

Burton et al. [13] analyze the data from motion capture, with an actor performing different tasks with different emotions, to introduce an automatic Laban quantification. They determine six different neutral hand pathways for the motion capture actor to perform with six basic emotions. They rate recorded clips by Laban Certified Movement Analyst. They quantify the captured motion data as kinetic energy, acceleration, velocity, and so on, at each time step, which would contribute to a Laban Shape or Effort parameter. They train a Hidden Markov model and use the Viterbi algorithm to generate the motion with desired emotion, based on a given motion path. They evaluate resulting animations with a user study and conclude that some emotions like sadness share common aspects with anger and fear, making it harder to distinguish based solely on body movement. However, results show that the distinction between opposite emotions is successfully perceived using their approach.

Masuda and Kato [33] introduce a method to add emotion to body movements of human form robots, using features based on Laban Movement Analysis. They introduce a mapping from emotions of pleasure, anger, sadness, and relaxation into movement features of space, time, weight, inclination, height, and area. Robot motion requires additional balancing; therefore, motion adjustments are limited. They make a user study on the perceived emotion of the robot to evaluate the introduced method.

Randhavane et al. [34] use expressive features of gait and gaze to express agent emotions. Their proposed architecture works in real-time to express neutral, happy, angry and sad emotions through a movement modification process. They focus on scenarios with multiple agents and aim to increase the sense of presence in VR. Based on a user study they introduce a data-driven mapping between gait features and perceived emotions.

Additionally, Shvo et al. [35] introduce an empirically-based interplay between personality, emotion, mood, and motivation of an agent. They use personality to influence the impact of emotions, which alter the mood of the agent and determine the importance of its motivations. Their model of agent motivation is based on the 16 Basic Desires Theory of Steven Reiss so that an agent would aim to fulfill these desires with varying importance levels. They use research from Dialectical Behavior Therapy to map emotions into motivations. They run a heuristic-based planner where agents could perform various actions based on their motivations in a scenario, and reflect on different actions each agent takes based on its personality and emotions. Instead of relying on the visual and vocal cues, they express personality through motivational behavior selection.

Our system uses an OCEAN-to-LMA parameter mapping similar to the one described in Durupinar et al. [7]. However, in terms of the application of the LMA parameters into the movement modifiers, our system uses a different approach. We calculate variable attraction weights for the end effectors towards the Laban shape quality anchors using the data obtained by preprocessing the base animations. Furthermore, we make use of vocal adjustments and additive facial expressions to enhance the expressed personality. Instead of focusing on multi-agent, heterogeneous crowd scenarios where there are agents of different personalities, we focus on conversational scenarios with one single agent interacting with the user.

Smith et al. [32] use movement modifications with the similar aim of expressing personality, however, they exclude the agent's face completely while we use the face as an enhancing element. Although they include speech in various experiments, they do not focus on the influence of vocal features on perceived personality. We consider the vocal features of the speech to be an effective personality cue. On the other hand, evaluation of their system is based on a two-dimensional space of plasticity and stability, while we focus on one-dimensional channels per OCEAN factor and leave the analysis of multi-dimensional OCEAN space as future research. In other words, when we adjust one OCEAN factor, we evaluate based on that particular factor, and yet to investigate how the remaining factors are influenced.

Burton et al. [13] and Masuda et al. [33] use LMA to express emotions, while we use LMA and emotions that emerge as facial expressions as a way of expressing personality. The correlation between emotions and personality could enhance the perception of both of them. The inclusion of emotions could help to express the perceived personality in a better way, as our results point out in Chapter 7. On the other hand, we expect personality to have a similar influence on the perception of emotions, based on the correlation between these two entities.

2.5.2 Relating Text and Speech to Personality and Emotions

Golbeck et al. [36] introduce a model to predict OCEAN personality based on people's profile information and writings on social media. They make a 45-question survey to users to analyze their personality and gather their social media data including relationship status, education history, gender, favorites, personal description and a collection of status update text. They find correlations between personality and various linguistic features including swear words, social processes, positive emotions, anxiety, work, and money-related words. Other features such as the number of friends and relationship status are also found influential on personality.

Mairesse et al. [37] use linguistic cues to predict OCEAN personality from conversation text. They compose a large set of language parameters from earlier works and train classification and regression models using a labeled text corpus. They also use set with voice data to incorporate speech features. They mostly focus on extraversion because it is more prominent in text and many other pieces of research investigate extraversion in text. Using the same model, they generate text with OCEAN personality in [38]. Generated lines comment about a restaurant with different extraversion scores, and a user study is made to evaluate perceived extraversion. Mairesse [5] perform a comprehensive study that takes into account all OCEAN dimensions, concluding extraversion is the easiest factor to express in text, and openness is the hardest.

Polzehl [4] focuses on speech-based personality assessment. Using the NEO Five-Factor Inventory (NEO-FFI) [39] questionnaire they label a speech dataset with OCEAN personality factors and extract sound features such as intensity, pitch, and loudness to train a Support Vector Machine. The resulting model performs best for extraversion, but for high agreeableness it is inconclusive. They record an inverse correlation between perception of neuroticism and extraversion. Their findings form a basis for our vocal feature adjustments.

Gilpin et al. [40] use Support Vector Machines and Hidden Markov Model to classify personality based on speech signals. They use the SSPNet Speaker Personality Corpus that contains speech clips labeled with OCEAN personality to train the model. They make a user study with new speech clips and ask users to rate them. The results of the trained model for these new speech clips show significant similarity to the users' personality assessment.

Charalambous et al. [41] introduce a procedural speech animation method that expresses emotions of happiness, sadness and anger. The model uses transcription of the speech to extract phonemes that are mapped into visemes. Audio of the speech is analyzed in terms of its pitch and intensity. High pitch and intensity are correlated with high arousal emotions such as joy and anger, and low pitch and intensity is related to low arousal emotions such as sadness and boredom. Audio features influence the articulation of phonemes to reflect emotions.

Chapter 3

The Proposed Architecture

We implement the system in Unity [42] using C#. It consists of two main components, as shown in Figure 3.1, one for handling the current scenario and the other for animation modification process. The user speech triggers the agent to respond with a dialogue speech and animation. We choose the corresponding dialogue unit according to the user intent and vocalize it based on the agent personality. We choose the corresponding base animation according to the scenario state and modify it based on the agent personality.

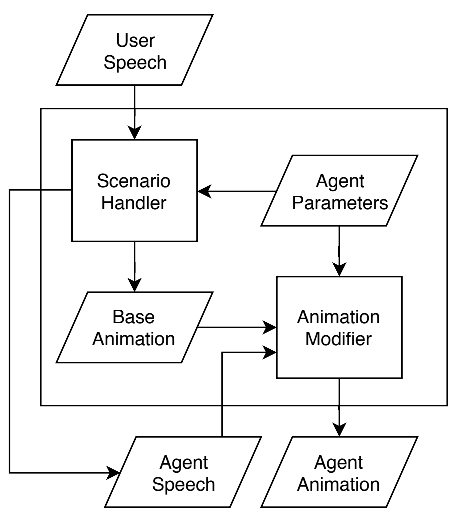


Figure 3.1: The interaction between the main components of the system.

Scenario Handler takes turns until it reaches the end state of the current scenario. Each turn starts with the user speaking into the microphone. We transcribe the user's speech using Watson Speech to Text API [43] in real-time. The recognized text appears on the screen, and with the user's "submit" command, we send the transcript to Watson Assistant, which is a pretrained unit responsible for finding the intent of the given text. The training of Watson Assistant requires setting entities and intents using multiple example text. For example, for "Show Passport" intent we train the system using examples in Table 3.1 so that it recognizes inputs in a similar direction such as "May I take a look at your passport?" and "Where is your passport?". Entities are trained in the same manner, and they are recognized within the intent.

Table 3.1: The training examples for Watson Assistant.

Intent: "Show Passport"
Can I see your passport?
Passport please.
Show me your passport, please.
May I check your passport?
Could you show me your passport?
Can I look at your passport?
Could you give me your passport, please?
May I have your passport?

We use the found intent and entities to determine the corresponding agent dialogue sentence and the accompanying base animation. First, we select a Dialogue Unit that contains different OCEAN Alternatives. Then, we select a suitable OCEAN Alternative as the Agent Dialogue Sentence based on OCEAN parameters of the agent. We use the state machine of the current scenario in this operation.

We switch the agent's current animation into the chosen base animation, which is then processed by the Animation Modifier in real-time. We convert the chosen dialogue sentence into speech using Watson Text to Speech API. We adjust the vocal features of the generated speech according to the OCEAN parameters of the agent. When the agent speech is ready it starts playing. Animation Modifier is responsible for animating the mouth shape to fit the speech in real-time.

Scenario Handler waits until the agent speech ends and then checks whether the end state of the current scenario is achieved. If the scenario is on an end state, this component finishes executing. If the end state of the scenario is not yet reached, then the state of the current scenario is updated and a new turn starts with the microphone listening to the user's speech. Figure 3.2 shows the Scenario Handler flowchart.

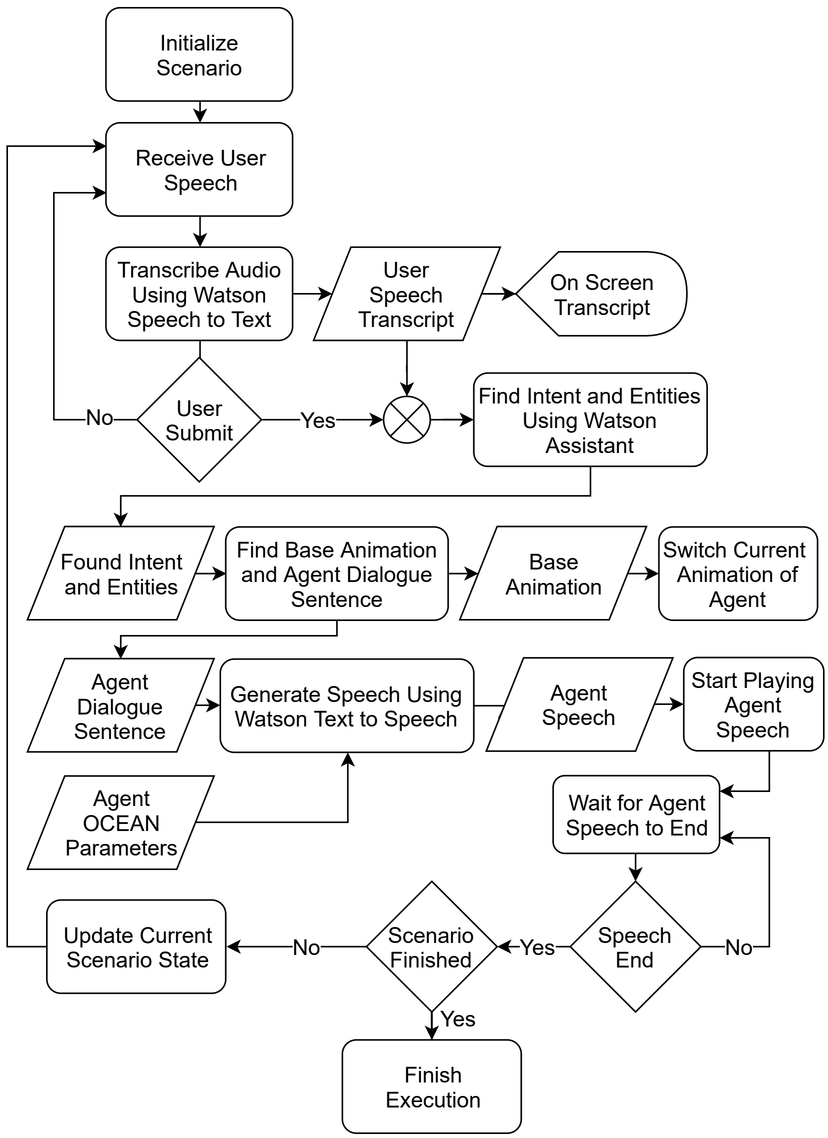


Figure 3.2: The Scenario Handler flowchart.

Animation Modifier works at each frame to alter the base animation into the final animation that fits the agent personality. The agent always plays an animation, and Scenario Handler is responsible for switching between animations. If the agent is talking, we analyze the speech audio to animate the mouth shape. We use Oculus Lip Sync [44] for extracting visemes and map them into facial shape key values. We determine emotions based on the chosen agent dialogue sentence at the beginning of the speech and blend them into the facial expression. Emotions decay with time and we set them to new values with each new speech. Figure 3.3 depicts the Animation Modifier flowchart.

We map the OCEAN parameters of the agent into the movement parameters: Laban Shape, Laban Effort, and Nonverbal Communication. Because we store these parameters per agent, we calculate them only when an agent's personality changes or at initialization.

We calculate Inverse Kinematics (IK) targets at each frame based on Laban Shape parameters. We use IK Targets for attracting hands towards anchor points related to Laban Shape Qualities. We base the attraction on current distances of hands to anchor points in time. We calculate the attraction weights by preprocessing animations. After this step, we further process the refined animation using additional rotations by Laban Effort and Nonverbal Communication parameters. We also interpolate the facial shape keys using the target values during this process. We show the resulting pose at each frame, producing the final animation. We use Adobe Fuse [45] for creating 3D human models with a skeletal rig and facial shape keys.

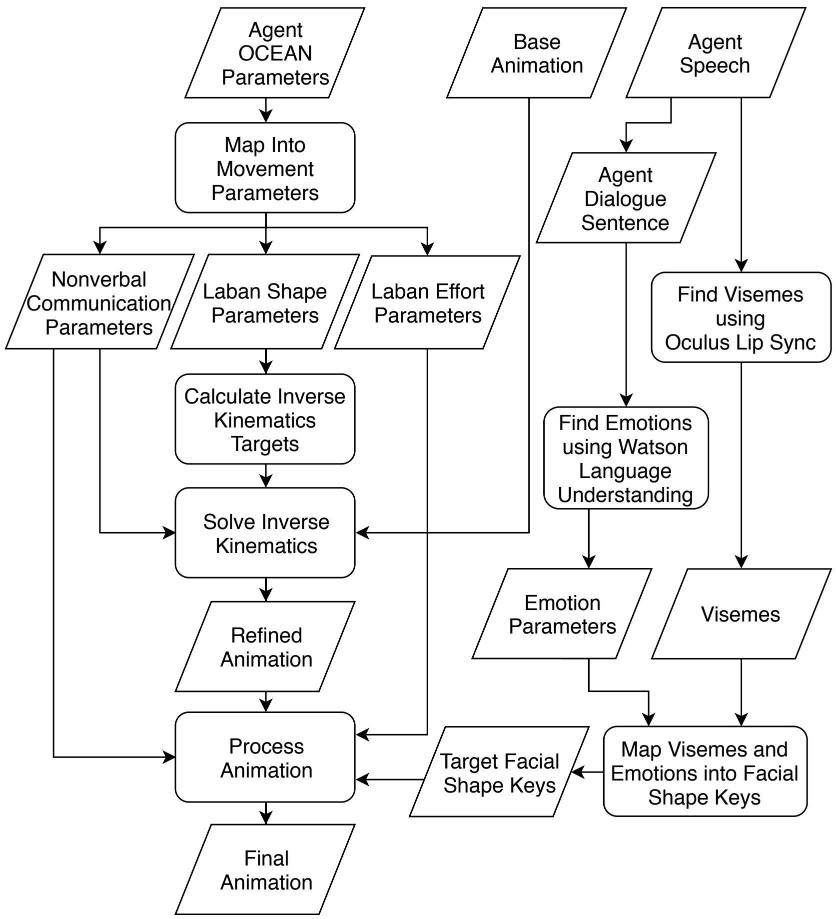


Figure 3.3: The Animation Modifier flowchart.

Chapter 4

Modifying the Base Animation

We modify the base animation of the agent according to multiple parameters in real-time. We perform various modifications by preprocessing the animation clips that will be used at the initialization step. The animation modification unit is responsible for producing a new set of bone rotations and facial shape key values, given the current bone rotations of the sampled animation clip, OCEAN personality parameters, and current speech of the agent. We group the modifications according to the theory in use, including Laban Shape Parameters, Laban Effort Parameters, and Additional Nonverbal Communication Cues.

4.1 Laban Shape Parameter Modifications

4.1.1 Calculating Attraction Weights

Laban Shape Qualities describe the change in the body according to time, in three different axes of the body: *longitudinal*, *frontal*, and *sagittal*. On the longitudinal axis, the movement could be rising or sinking, on the frontal axis, it could be spreading or enclosing, and on the sagittal axis, it could be advancing or retreating.

We use three parameters: up, side, and forward for longitudinal, frontal and sagittal axes, respectively; positive values indicating the direction that the parameter is named with. Each parameter could take a floating value between -1 and 1, where a zero value means no changes for that axis. Intermediate values determine the magnitude of attraction towards a direction.

Ultimately, we would like to orient the end effectors towards a direction. We place six anchor points for each hand, originating from the center of the body, taking one unit towards up, down, left, right, front, and back (see Figure 4.1). We calculate the distance between each hand and each anchor point in the preprocessing stage. The distance ratio between the anchor points of the opposite ends determines the attraction value for that axis. After the modification, this value should approach to the desired one, without breaking the essence of the base animation.

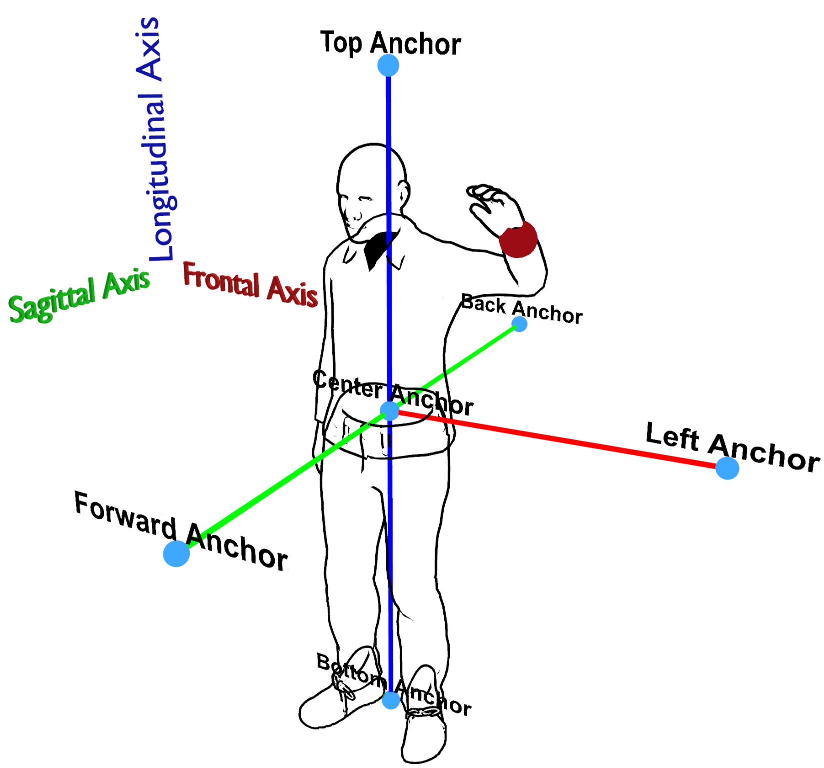


Figure 4.1: The anchor points for the left hand.

Consider a pointing motion where the hand starts at a resting position and ends in a forward-pointing state. If we would like to increase the advancing aspect of this animation, the hand should not start in a forward position. Instead, we would like to push the hand further away from the body while it is pointing. In other words, we would like to use the highest attraction weight towards the front anchor when the hand is closer to the anchor point and the lowest one when it is farther from the anchor point. Figure 4.2 shows a waving motion with constant and varying attraction weights.

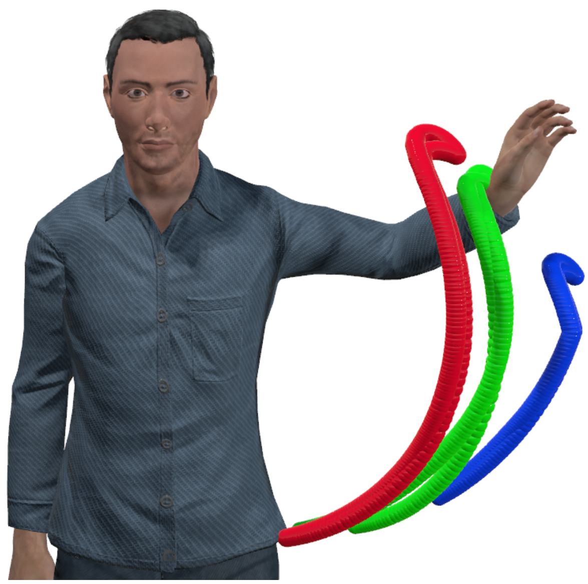


Figure 4.2: The base animation and its modified versions for a waving motion. The base animation is shown in red. The side anchor is active in modified animations. The animation with a constant anchor weight is shown in blue and the one with variable anchor weight is shown in green.

As another example, consider an idle animation where hands move slightly around their resting position. In case we would like to increase the advancing aspect, hands should not move towards the front anchor dramatically. To prevent this, we calculate an attraction factor for each anchor based on closest-farthest hand distance difference over the maximum possible difference of hand distances. This factor is always between 0 and 1, decreasing the magnitude of attraction when it is necessary. The attraction weight is the current distance mapped into the range [0, 1], multiplied by the attraction factor:

$$F_{ab} = \frac{|min(\mathbf{d}_{ab}) - max(\mathbf{d}_{ab})|}{|DMin_{ab} - DMax_{ab}|} ,$$
$$W_{ab}(t) = F_{ab}\left(1 - \frac{(\mathbf{d}_{ab}(t) - min(\mathbf{d}_{ab})}{max(\mathbf{d}_{ab}) - min(\mathbf{d}_{ab})}\right) , \tag{4.1}$$

where $W_{ab}(t)$ is the attraction weight of bone b to anchor a at time t, F_{ba} is the attraction factor of anchor a for bone b, $\mathbf{d}_{ab}$ is the vector of per frame distance from anchor a to bone b, $DMin_{ab}$ is the closest possible distance between anchor a and bone b, and $DMax_{ab}$ is the farthest possible distance between anchor a and bone b. $DMin_{ab}$ is found by using the limb length from the root of the bone chain. Because we use this value for Hand IK, we project a line from the shoulder position towards anchor a and calculate the distance as $|Pos_a - Pos_b|$. $DMax_{ab}$ is found similarly.

4.1.2 Calculating Inverse Kinematics Targets

We calculate hand IK targets at each frame using the attraction weight of each anchor, as well as Laban Shape Quality parameters, as described in Algorithm 1. We calculate Laban Shape Quality parameters (LSQ_x, where $x \in \{up, left, forward\}$) based on OCEAN parameters, and they are constant during an animation.

Each LSQ_x determines the attraction towards a specific anchor. The attraction is towards a specific direction of the corresponding axis based on the sign of LSQ_x, and with a factor of $(|LSQ_x| \times W_{ab}(t))$. We use linear interpolation for finding IK Target positions per LSQ_x, and calculate the average as the final IK target position for bone b. We perform this calculation for the left- and right-hand bones. We then feed the calculated positions into the IK Solver to generate a refined pose.

We use the internal IK Solver of Unity. In the main loop, before IK Update Pass, we calculate Target IK Positions for both hands and passed them into the IK Solver. We set the IK weight to one because we already used the attraction weight in the Target IK position calculation. We further modify the new pose

Result: The position of IK Target for Bone b at Time t
$IKPos_b$ = Zero;
if $LSQ_{up} > 0$ **then**
  $IKPos_b += Lerp(Pos_b(t), Top\ Anchor, LSQ_{up} \times W_{up\text{-}b}(t))$;
else
  $IKPos_b += Lerp(Pos_b(t), Bottom\ Anchor, -LSQ_{up} \times W_{down\text{-}b}(t))$;
end
if $LSQ_{left} > 0$ **then**
  $IKPos_b += Lerp(Pos_b(t), Left\ Anchor, LSQ_{left} \times W_{left\text{-}b}(t))$;
else
  $IKPos_b += Lerp(Pos_b(t), Right\ Anchor, -LSQ_{left} \times W_{right\text{-}b}(t))$;
end
if $LSQ_{forward} > 0$ **then**
  $IKPos_b += Lerp(Pos_b(t), Front\ Anchor, LSQ_{forward} \times W_{front\text{-}b}(t))$;
else
  $IKPos_b += Lerp(Pos_b(t), Back\ Anchor, -LSQ_{forward} \times W_{back\text{-}b}(t))$;
end
$IKPos_b$ /= 3;

Algorithm 1: Calculating IK target positions for hands.

produced by the IK Update Pass to generate the final animation.

4.2 Laban Effort Parameter Modifications

4.2.1 Blending Bone Rotations

We map Laban Effort (LE) parameters into individual bone rotations that are blended into the current pose. LSQ parameter modifications handle the time-based impact of weight and space components. The bone rotation blending process adds the constant impact of Laban Weight and Space components.

The Space component attributes interest in the environment. Blended rotations for this component are shown in Table 4.1. Positive space aims an extended open posture and negative space aims a contracted close posture. We determine the maximum and minimum rotations manually in a way that looks natural. The blending factor is determined by the magnitude of the space parameter.

Table 4.1: The Space Effort rotation mapping.

Bone and Rotation Axis	**Space -**	**Space +**
Shoulder - Longitudinal	Towards Front	Towards Back
Shoulder - Saggital	Towards Center	Towards Side
Upper Arm - Saggital	Towards Center	Towards Side
Foot - Longitudinal	Towards Front	Towards Back
Hand - Longitudinal	Towards Back	Towards Front

The weight component attributes the gravitational force on the body. Blended rotations for this component are shown in Table 4.2. A positive weight aims a lower posture whereas a negative aims a higher posture.

Table 4.2: The Weight Effort rotation mapping.

Bone and Rotation Axis	**Weight -**	**Weight +**
Spine - Frontal	Towards Up	Towards Down
Neck - Frontal	Towards Up	Towards Down
Shoulder - Saggital	Towards Side	Towards Center
Upper Arm - Frontal	Towards Up	Towards Down

4.2.2 Relative Change of Animation Speed

As an example, consider a pointing gesture that starts and ends with the same idle pose. The hand would translate into the pointing position and stay still for a limited time before returning to the idle pose. Basic speed up for such animation would shorten the pointing time, and in extreme cases decrease the realism instead of enhancing sudden time. The preferred result is to make the translational motion quicker, without disturbing the pointing duration dramatically.

We use an approach based on limb speeds, and in particular speed of each hand. The animation is preprocessed with a constant sample rate to calculate hand speeds per step, following Algorithm 2 Each step is ranked according to the average hand speed and mapped to a speed factor. The speed factor is higher for the time step with the fastest hand speed and lower for the slowest. The speed factor boundaries are adjustable.

for *Animation A in AnimationSet* **do**
  $FrameCount = A.Duration \times SamplesPerSecond$;
  $GapCount = FrameCount - 1$;
  $NormalizedFrameDuration = 1/FrameCount$;
  Initialize Array $Speed[GapCount]$;
  for *Integer T = 1 to FrameCount* **do**
    Sample A at Time $= (T-1) \times NormalizedFrameDuration$;
    Vector V_{t-1} = Bones.Hand.Position;
    Sample A at Time $= T \times NormalizedFrameDuration$;
    Vector V_t = Bones.Hand.Position;
    $Speed[T] = |V_{t-1} - V_t|$
  end
  Initialize Array $Rank[GapCount] = 1 : FrameCount$;
  $Sort(keys = Speed, items = Rank)$;
end

Algorithm 2: Preprocessing for the animation speed.

During runtime, looking at the normalized time of animation, the corresponding time step is found following Algorithm 3; and the animation speed, which is an internal Unity parameter that determines current animation's play speed, is adjusted based on the preprocessed speed factor and the time effort parameter. The resulting animation appears faster or slower, without breaking relative stops.

Result: Modified Speed at Current Gap
T = Normalized Time of Current Animation;
CurrentGap = $T \times GapCount$;
CurrentRank = $Rank[CurrentGap]$;
NewSpeed = Map CurrentRank from [1:*FrameCount*] to [MinSpeed:MaxSpeed];
Return NewSpeed;

Algorithm 3: Setting the animation speed at each time frame.

4.2.3 Adding Fluctuation using Circular Walk on Perlin Noise

LE_{FLOW} determines the fluctuation of limbs. We use controlled random values for additive bone rotations per axis. Minimum and maximum limits are set for each rotation by multiplying absolute limits with LE_{FLOW} scaled to the range

$[0, 1]$. To generate repetitive random values, we use circular walks on Perlin Noise.

We initialize a circular walker by setting a random center point in the range $(0, 1)$ for each dimension of 2D space. We generate another random value in the range $(-2\Pi, 2\Pi)$ as the starting angle and set a radius value. At any given time, we calculate the position of the circular walker as $(Center_x + Cos(angle) \times Radius, Center_y + Sin(angle) \times Radius)$. At each frame, we update the angle by a constant *deltaAngle*. We calculate the Perlin value at the position of the walker as the fluctuate angle. Due to the nature of the Perlin noise, blended rotations appear smoothly during an animation.

We set a timer for each walker. The walker changes its orbit when the timer reaches zero. We calculate the new center of the circle from the current walker position with a new random angle and radius; thus, the walker orbits a different center without jumping to a different position.

Grayscale textures could be used for sampling, instead of the Perlin noise. The current walker point would be sampled from the corresponding texture position, so to prevent jumps the texture should be seamless and smooth. Certain textures could introduce different styles of fluctuation, and it can also be used as a way to generate idle animation.

4.3 Additional Nonverbal Communication Cue Modifications

Facial expression of the agent is a blend of six different expressions that show anger, sadness, happiness, disgust, surprise, and neutrality. There are some kinesics that are not covered by Laban Shape Quality and Effort modifications directly. We include them as additional nonverbal communication cues and adjust the target values for facial shape keys of the agent accordingly (cf. Table 4.3):

Each facial expression has a value between 0 and 1. Each expression decays as time passes. Expressions could be adjusted incrementally. For example, a value could be added to happiness at different time steps. We use this approach in mapping the agent dialogue into emotions. At the beginning of each dialogue unit, the agent's facial expression is adjusted using additive values. We determine the eye blink rate of the agent by a factor related to neuroticism. For example, a neurotic agent would blink frequently [46].

Table 4.3: Additional nonverbal communication cues and their corresponding shape key values. Because neutrality doesn't effect any shape keys, we did not include it in the table.

Target	**Anger**	**Disgust**	**Sadness**	**Happiness**	**Surprise**
brows down	1.0	0.5			
cheek puff	0.02				
frown	0.4	0.1	0.8		
mouth down	0.1	0.1	0.1		
mouth narrow	0.2				0.2
squint	0.3	0.3	0.1	0.4	
brows up			0.1	0.3	1.0
eyes wide					0.8
mouth open				0.1	0.3
smile				1.0	
brows outer lower			1.0		0.5
brows in			0.1		
jaw backward			0.1		
nose scrunch		0.8			
mouth up		0.1			
jaw forward		0.1			
upper lip out		0.5			
lower lip in		0.2			
mid mouth		0.2			

Chapter 5

OCEAN Personality and Emotion Mapping

5.1 Mapping Personality to Movement Parameters

We are able to calculate the movement parameters from OCEAN parameters mathematically, as they both use floating point values between -1 and 1. Movement parameters are grouped as *Laban Effort (LE)*, *Laban Shape Quality (LSQ)*, and *Nonverbal Communication (NC)* parameters. We consider some OCEAN parameters to have more impact on certain movement parameters. Thus, we use different weights in the mapping, which are determined experimentally (see Equations 5.1, 5.2, and 5.3). More effective OCEAN parameters have a coefficient of 1.5 and less effective ones have a coefficient of 0.5.

We base our OCEAN-to-Laban-Effort mapping on [7]. Laban Effort parameters are calculated as follows:

$$
\begin{aligned}
LE_{space} &= (E + O + 0.5(N - C))/3 \\
LE_{weight} &= (A + 0.5(O + E))/2 \\
LE_{time} &= (E + N - 1.5C - 0.5A)/4 \\
LE_{flow} &= (1.5N + 0.5E - C + O)/4
\end{aligned}
\tag{5.1}
$$

where O, C, E, A, N denotes Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, respectively; and LE_x denotes the corresponding Laban Effort parameter. In the mapping Durupınar et al. [7] introduce, all related OCEAN parameters and LE_x use the same weight.

We calculate the Laban Shape Quality parameters as follows:

$$
\begin{aligned}
LSQ_{up} &= (O + C + E + A)/4 \\
LSQ_{side} &= (1.5O + E)/2.5 \\
LSQ_{forward} &= (E + N)/2
\end{aligned}
\tag{5.2}
$$

where LSQ_x the corresponding Laban Shape Quality parameter.

We use Nonverbal Communication parameters for modifications that are not covered directly by LMA. We introduce the following mapping from OCEAN to Nonverbal Communication parameters. The equations are formed experimentally, based on [19]:

$$
\begin{aligned}
NC_{SpineBend} &= -(E + 0.5A)/1.5 \\
NC_{NeckBend} &= -(0.5O + 0.5C + E + 0.5A)/2.5 \\
NC_{Sink} &= -(O + 0.5C + 0.5E)/2 \\
NC_{FingerBend} &= -(O + A - N)/3 \\
NC_{BlinkGap} &= (C - N)/2 \\
NC_{BlinkSpeed} &= (C - N)/2 \\
NC_{Expresiveness} &= E \\
NC_{EmotionDecay} &= (N - E)/2
\end{aligned}
\tag{5.3}
$$

5.2 Personality-based Response in Dialogue

We use the term Dialogue Unit (DU) to indicate the text agent could speak during its turn in a dialogue. Each DU includes the original text, as well as text for 10 possible OCEAN Alternative (OA).

Although one line of dialogue could express each OCEAN factor to some degree, we assume one dimension is emphasized. Therefore, for one DU we store 10 different OA, each representing one positive or negative OCEAN personality factor. In our implementation, OAs are handcrafted following Mairesse's personality-text findings [5]. One example DU and accompanying OAs are visible in Table 5.1.

Table 5.1: Example OCEAN Alternatives for a Dialogue Unit from the Passport Scenario.

Type	**Text**
DU	"Ok, I will buy my return ticket."
$O(+)$	"Of course, I'm going to buy it for sure. I should have done it before; this will be a lesson for me."
$O(-)$	"Oh, ok. I will buy it."
$C(+)$	"Yes, I am going to buy my return ticket as soon as possible."
$C(-)$	"I will buy it... I will, as soon as possible."
$E(+)$	"Yes, I will buy the return ticket immediately. Thank you, officer."
$E(-)$	"Ok, I will."
$A(+)$	"Thank you very much for reminding, I'm going to buy it as soon as possible."
$A(-)$	"Well, I have to, you know. I will buy it."
$N(+)$	"I. . . I am going to buy it. I will buy it as soon as possible."
$N(-)$	"Sure, I will buy it as soon as possible."

The state machine of the scenario is responsible for choosing an appropriate DU for the agent to talk, given the user's intent. According to the agent's OCEAN parameters, we probabilistically select an OA.

$$\begin{aligned} P(OA_X{}^{+}) &= max(0, min(X, 1)) \text{ and} \\ P(OA_X{}^{-}) &= max(0, min(-X, 1)), \end{aligned} \tag{5.4}$$

where X is the corresponding OCEAN parameter, $X \in \{O, C, E, A, N\}$, ${OA_X}^+$ is the positive OCEAN alternative of parameter X, and ${OA_X}^-$ is the negative OCEAN Alternative of parameter X.

Using Equation 5.4, we calculate a probability for each OA. We clamp the values to the range [0, 1] so if the OCEAN factor is negative, the probability of corresponding positive OA is set to zero. This means an agent with negative extraversion, for example, will never use an OA of positive extraversion. As another example, if $E = 0.8$ and $A = 0.4$, then the agent will be more likely to choose the E^+ alternative. The sum of all OA is scaled to 1 and one OA is selected randomly to become the agent's response.

5.3 Changing Speech Features to Support Personality

The vocal features of the speech are adjusted according to [4] to reinforce personality. Watson Text to Speech API [43] supports multiple transformable voices. We use "en_US_Michael" for male and "en_US_Allison" for female voices. We use the following parameters that could take values between -100 and 100 for transforming the voice:

Pitch: The frequency of the voice; higher values correspond to higher frequencies. Speech sounds more excited when the pitch is high, and calmer when it is low.

Pitch range: The variation in pitch during a speech; higher values have more pitch variation. Higher pitch range sounds more melodic and dynamic, where lower pitch range is perceived monotonous.

Rate: The speed of the voice; higher values correspond to faster speech. A higher rate is perceived as hurried and excited, on the other hand, a lower rate sounds calm.

Breathiness: The amount of escaping air during sound production; higher values correspond to breathier voice. Higher breathiness is perceived calmer.

Glottal tension: The hardness of the voice, higher values sound harder. Higher glottal tension is perceived dynamic and tense, while lower sounds calmer and softer.

We use the mappings in Equation 5.5 to calculate each vocal feature:

$$\begin{aligned}
&\mathit{Pitch} = 80(O + 0.5(C + E - A - N)) \\
&\mathit{PitchRange} = 50(O + E) \\
&\mathit{Rate} = -70(O + C - E + A)/4 \\
&\mathit{Breathiness} = 25(A + O) \\
&\mathit{GlottalTension} = -100(O - C + A)/3
\end{aligned} \tag{5.5}$$

It is meaningful for the current emotion of the agent to influence vocal features as well, however, we did not include it in our implementation.

5.4 Finding and Mapping Emotions

Expressing emotions using facial features has an important role in increasing the realism of human animation. It is possible to assign different emotions based on time or events; however, it is more feasible to use dialogue content to predict emotions in a conversational agent. We use Watson Natural Language Understanding API [43] to estimate emotions of joy, anger, disgust, sadness, and surprise from the agent's current line of text. Each emotion has a value between 0 and 1.

When the agent starts speaking the current line of text, the emotion values of the corresponding text are used to set the agent's internal facial expression parameters: happiness, anger, disgust, sadness, and surprise. On top of that, using additive facial expression values introduced in Facial Expression - Personality Experiment in Section 7, we add constant expression values based on personality.

We calculate the internal emotion decay rate of the agent using neuroticism. The agent's emotion parameters decay at each frame based on this rate. We then map the emotion parameters into the target shape key values that control the facial expression. The model generated from Adobe Fuse [45] has 50 shape keys that control a specific part of the face. We use the mapping between different facial shape keys and facial expression parameters that are introduced in Additional Nonverbal Communication Cue Modifications.

The agent stores arrays of current and target values for shape keys. We update the current value of each shape key towards its target at each frame and store the update amount per shape key in a separate array. In the current implementation, we alter only the blink speed by personality, and the other speeds remain constant; however, it could be more realistic to adjust the animation speeds of other shape keys as well, according to personality or emotion.

We update the 3D mesh of the agent using current shape key values multiplied by an expressiveness factor, that is determined by the extraversion parameter. This makes it possible to have subtle expressions on introvert agents and exaggerated ones on extraverts.

Chapter 6

Scenarios

We implemented different scenarios to test the conversational agent. Each scenario includes an appropriate scene setup, one main agent and a dialogue state machine. The scenarios aim to direct the user input in a specific direction because the implemented agent is not designed for general dialogue. Although we implemented the scenes using ALICE AIML (Artificial Intelligence Markup Language) and the agent is capable of answering general questions, we do not bring this stage into the user study. This is because ALICE is not designed for long chains of dialogue. Because it has a sizable answer set, it is not feasible to handcraft dialogue alternatives for each answer, fitting different personality types. The user is either given a task at the beginning of a scenario or guided with state-specific descriptions inside the scene. Scenario-specific data per agent is generated if necessary, such as name, age, and gender.

We implemented *Introduction*, *Fastfood*, and *Passport Officer* Scenarios, which we explain in the sequel.

6.1 Introduction Scenario

In this scenario, the user is asked to learn certain information about the agent including name, age, occupation, and the city agent lives in. Neutral outside scene setup is used for this scenario, and the upper body of the agent is visible to the user. Figure 6.1 shows still frames from the scenario. In this scenario, the dialogue state machine supports non-linear flows, as seen in Figure 6.2, yet the user could tend to follow a certain order. The user can end the conversation with a farewell at any time.

Figure 6.1: Still frames from Introduction Scenario. A low conscientiousness agent (careless, negligent) is used. White subtitles correspond to agent's lines.

The name, age, occupation, and city of the agent is randomized at each execution. OCEAN Alternatives of agent Dialogue Units have a minimal distinction in this scenario because the questions of the user are rather direct. An example Dialogue Unit with its OCEAN Alternatives is given in Table 6.1.

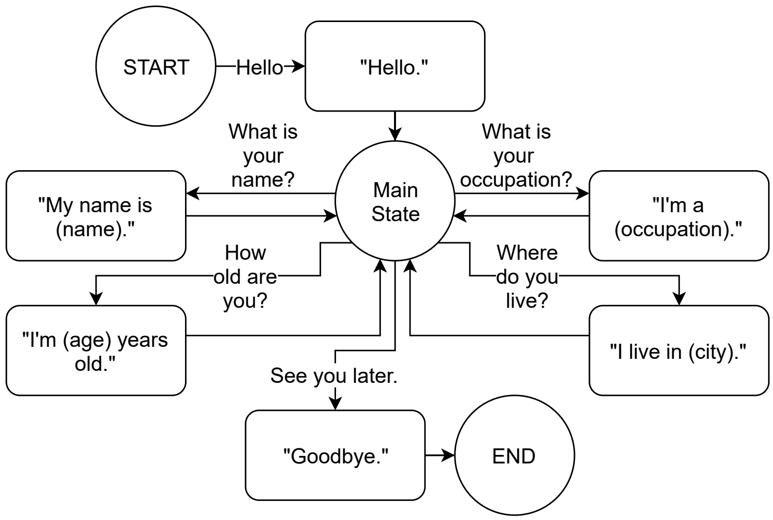


Figure 6.2: The dialogue state machine of Introduction Scenario.

Table 6.1: Example OCEAN Alternatives for a Dialogue Unit from Introduction Scenario.

Type	**Text**
DU	“My name is (name).”
$O(+)$	“I am known as (name).”
$O(-)$	“I'm (name).”
$C(+)$	“My name is (name).”
$C(-)$	“Oh, well... My name is... (name).”
$E(+)$	“I'm (name), my friend.”
$E(-)$	“(name).”
$A(+)$	“My name is (name), nice to meet you.”
$A(-)$	“Why do you ask? It's (name).”
$N(+)$	‘Um... I... I am (name).”
$N(-)$	“My name is (name).”

6.2 Fastfood Restaurant Scenario

In this scenario, the user takes the role of a customer. The aim of this scenario is for the user to order some food by talking to the cashier agent, in a fast-food restaurant setup. We do not give specific information about what to order to the user. We expect the user to ask what food restaurant serves before ordering. Some products could be out of stock in this scenario. The cashier asks whether the user wants a menu or the big selection where it is applicable. At the end of the dialogue, the user has the option to pay with a credit card or cash. The scenario ends with cashier preparing the order. Figure 6.3 shows still frames from the scenario and Figure 6.4 gives the dialogue state machine of this scenario.

Figure 6.3: Still frames from Fastfood Scenario. An extravert agent (talkative, sociable) is used. White subtitles correspond to agent's lines.

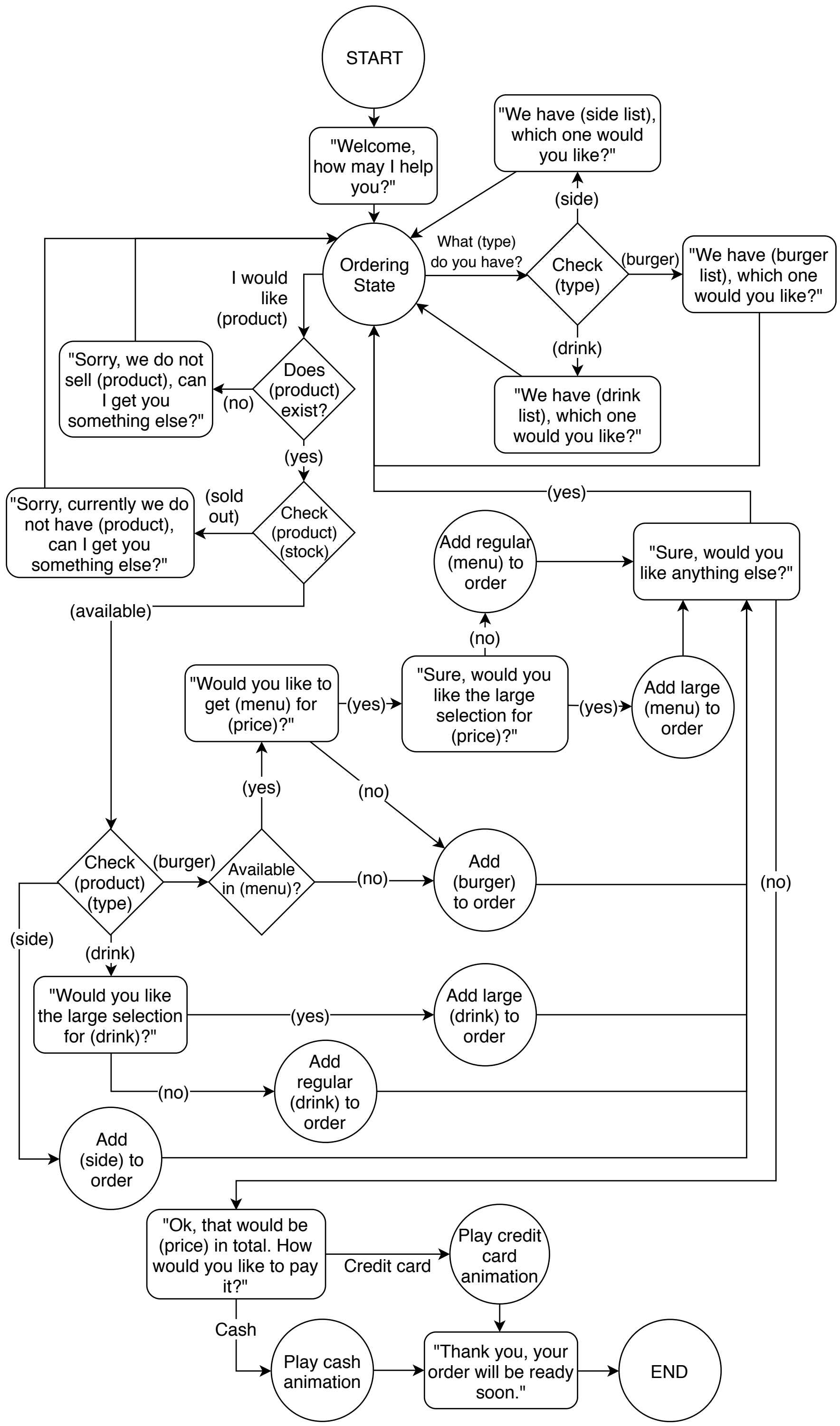


Figure 6.4: The dialogue state machine of Fastfood Scenario.

6.3 Passport Officer Scenario.

The user takes the role of a passport officer. This scenario aims to question the visitor agent. An airport setup is used for this scenario, as seen in Figure 6.5. The agent's passport information includes visa and passport issue and expiration dates, and a visa type. To enter the country, issue and expiration dates should be valid and the passenger's purpose of visit should be appropriate to the visa type. We show a guiding message on the screen according to the current state to help the user assume the role, but the user does not have to follow the guide and could decide on the final decision independently. We guide the user to ask the occupation and return ticket of the visitor as well. Figure 6.6 shows the dialogue state machine of this scenario.

Figure 6.5: Still frames from Passport Scenario. An intravert agent (quiet, reserved) is used. White subtitles correspond to agent's lines.

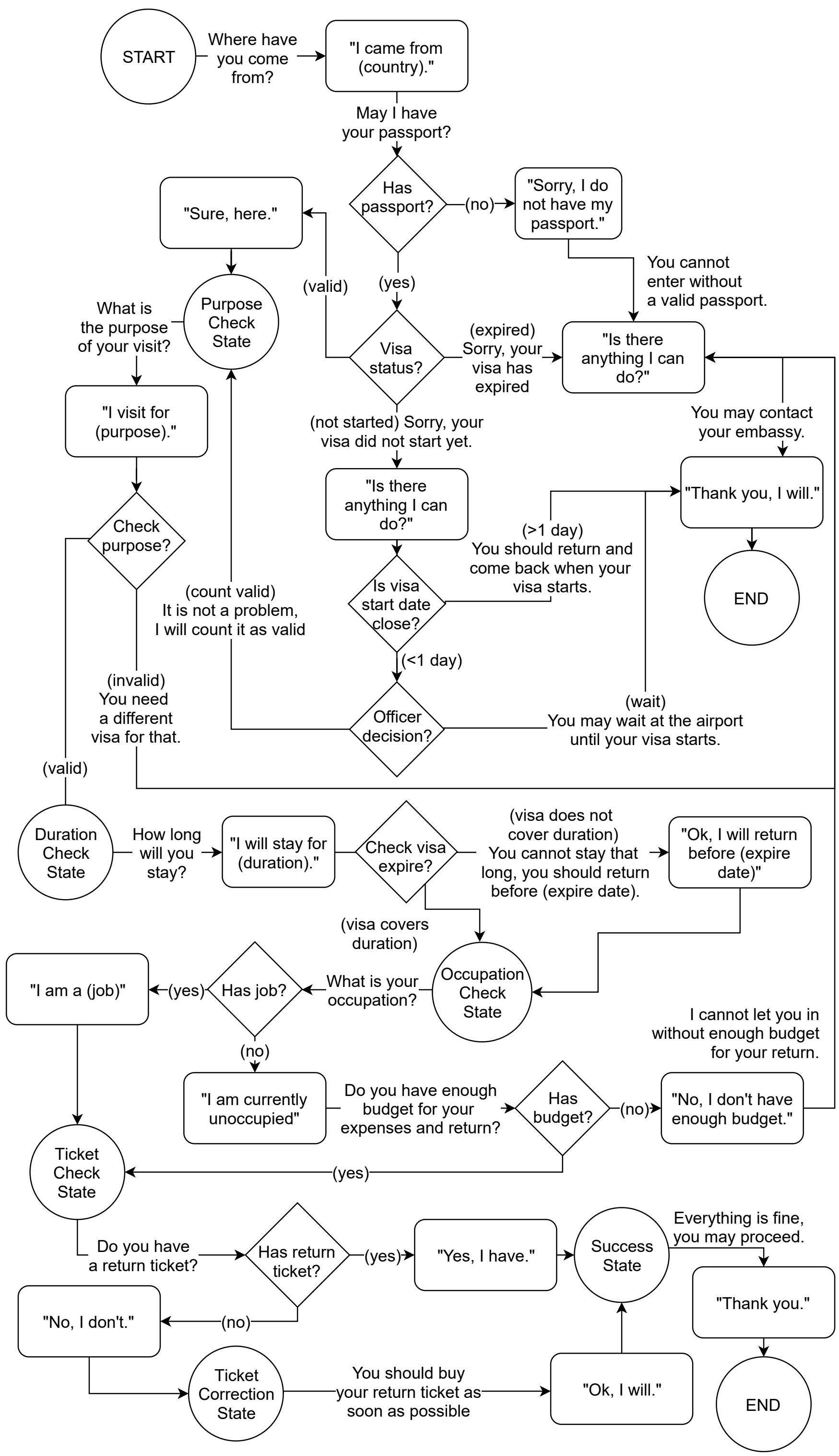


Figure 6.6: The dialogue state machine of Passport Scenario.

Chapter 7

Evaluation and Results

We performed user studies on Amazon Mechanical Turk [47] to evaluate the implemented system. To be able to perceive the influence of each module, we prepared five models where different modules are active. The users rated personality statements about a set of agent samples using a model, and OCEAN personality scores are calculated per sample. There were ten different samples in a set, each representing one OCEAN extreme. Difference between the score of an OCEAN factor for the positive and negative samples shows how distinguishable that factor is, using the corresponding model.

We use two experiments for tuning the system parameters; they are examined in the Initial Experiments section. They include choosing a neutral agent model and examining the relation between facial expressions and personality. We discuss the experiments that evaluate the five models in the Main Experiments section.

In every experiment, we measure the perceived personality of an agent using the Ten-Item Personality Inventory (TIPI) [48]. We show an image or a video to the user in each sample, including transcription of the conversation when applicable. Every main experiment includes a set of ten different samples, each representing a positive or negative extreme of an OCEAN Factor. Initial experiments have different numbers of samples, discussed in the related section. When a user participates in an experiment, he or she has to rate all samples of that experiment. The samples are presented to the user in random order. We did not allow for a user to participate in multiple experiments to prevent the bias that is caused by learning the agent.

For each sample, the user is asked to rate ten different statements based on the agent's personality. These statements are shown in Table 7.1. Each statement is related to one OCEAN factor positively or negatively. For example, if the user agrees that the agent looks extraverted and enthusiastic, this contributes to positive extraversion; on the other hand, disagreeing that the agent looks reserved and quiet would contribute to positive extraversion. The ordering of the statements are based on [48], and they remain the same for every sample and experiment.

Table 7.1: The statements used in OCEAN personality measure in each experiment. Each statement is related to an OCEAN factor in a signed manner. The value of the user's choice is determined by the mapping in Table 7.2, based on the sign and the personality factor. The values of the positive and negative factor are averaged and normalized to calculate the corresponding OCEAN factor score.

Statement	**Factor**
1. This character looks extraverted and enthusiastic.	E (+)
2. This character looks critical and quarrelsome.	A (−)
3. This character looks dependable and self-disciplined.	C (+)
4. This character looks anxious and easily upset.	N (+)
5. This character looks open to new experiences and complex.	O (+)
6. This character looks reserved and quiet.	E (−)
7. This character looks sympathetic and warm.	A (+)
8. This character looks disorganized and careless.	C (−)
9. This character looks calm and emotionally stable.	N (−)
10. This character looks conventional and uncreative.	O (−)

In each experiment, the user chooses one option per statement. The options are presented in a 7-Point Likert Scale and remain the same for every statement, sample, and experiment. Each user choice is converted to a value, following the mapping given in Table 7.2, based on the sign and the related OCEAN factor. For the OCEAN factor X, score is calculated as $\mathit{Normalize}(\mathit{Value}(X(+)) + \mathit{Value}(X(-)))/2$.

Table 7.2: The options given for each statement in Table 7.1 and the corresponding values for positive and negative OCEAN factors. We average and normalize the corresponding positive and negative values to calculate the score for each OCEAN factor.

Option	**Value (+)**	**Value (−)**
Disagree strongly	-3	3
Disagree moderately	-2	2
Disagree a little	-1	1
Neither agree nor disagree	0	0
Agree a little	1	-1
Agree moderately	2	-2
Agree strongly	3	-3

We normalize all OCEAN scores in the results. The values of 0.0, 0.5, and 1.0 correspond to negative, neutral, and positive personality factors, respectively. For every sample in an experiment, we calculate an OCEAN score per user, as in Table 7.3.

Table 7.3: The statement ratings given by one user for one sample in an experiment. For each OCEAN factor, there are two different statements: one positive and one negative. We convert the choices of the user into values using the mapping given in Table 7.2. We average and normalize the value of each statement pair to calculate the corresponding OCEAN factor score.

Factor	**Statement**	**Value**	**Calculation**	**Score**
O	O(+)	2	Normalize($(O_+ + O_-)/2$)	0.833
	O(−)	2		
C	C(+)	0	Normalize($(C_+ + C_-)/2$)	0.666
	C(−)	2		
E	E(+)	3	Normalize($(E_+ + E_-)/2$)	1.000
	E(−)	3		
A	A(+)	2	Normalize($(A_+ + A_-)/2$)	0.916
	A(−)	3		
N	N(+)	-2	Normalize($(N_+ + N_-)/2$	0.333
	N(−)	0		

In the main experiments, one user rates ten different samples, and the set of ten samples is rated by 50 different users. The OCEAN scores of every sample are analyzed using statistical methods. Although we calculate a score for all OCEAN factors of a sample, only the OCEAN factor of interest is used in the evaluation.

Because each sample is related to one positive and one negative OCEAN factor, we compare the scores of the factor of interest in the positive and negative samples. For example, Table 7.4 contains OCEAN scores per sample in an experiment, by one user. The openness difference of O(+) and O(−) is used as a metric to analyze the efficacy of the model in distinguishing openness. The scores for conscientiousness, extraversion, agreeableness, and neuroticism are not used in these two samples. For every factor of interest, we expect the mean score of the positive sample to be higher than the mean score of the negative sample.

Table 7.4: The ratings given by one user to a set of agent samples in one experiment. Each row provides calculations for one OCEAN factor of interest (either positive or negative). The difference indicates how much positive and negative samples are distinguished.

Sample	**O**	**C**	**E**	**A**	**N**	**Factor**	**Difference**
O(+)	0.833	0.667	1.000	0.917	0.333	O	0.500
O(−)	0.333	0.250	0.417	0.250	0.667		
C(+)	0.417	0.750	0.333	0.500	0.167	C	0.583
C(−)	0.500	0.167	0.250	0.500	0.750		
E(+)	0.417	0.667	0.250	0.417	0.333	E	0.083
E(−)	0.500	0.417	0.167	0.250	0.583		
A(+)	0.333	0.667	0.333	0.417	0.417	A	0.083
A(−)	0.333	0.500	0.250	0.333	0.500		
N(+)	0.333	0.333	0.417	0.500	0.667	N	0.333
N(−)	0.500	0.750	0.583	0.750	0.333		

To analyze the difference between two sets of OCEAN scores given to two different samples, we use two-tailed Welch's T-Test. We compare means of the same OCEAN factor of two opposite samples only, potential correlations between different OCEAN factors are out of the scope of this research. When there is a significant difference between the two sets of OCEAN factor scores, we also use mean differences as a comparison metric. To compare the difference between different models, we use one-way ANalysis Of VAriance (ANOVA) with Tukey Honestly Significant Difference (HSD), and report significant differences [49, 50].

Figure 7.1 illustrates ten different samples within a model experiment. In the example, we compare the openness means of O(+) and O(−) samples. We use Welch's T-Test for the comparison. For every OCEAN factor, we compare sample pairs similarly. To compare two different models, we calculate a set of mean differences per sample pair for each OCEAN factor and compare these mean differences from different models. For example, we compare the openness mean difference from Base Model and Model V (see Figure 7.1). We use One-way ANOVA with Tukey HSD for this comparison; we apply it on each possible model pair for each OCEAN factor. The openness difference of Base Model and Model V corresponds to a value in one cell in Table 7.16.

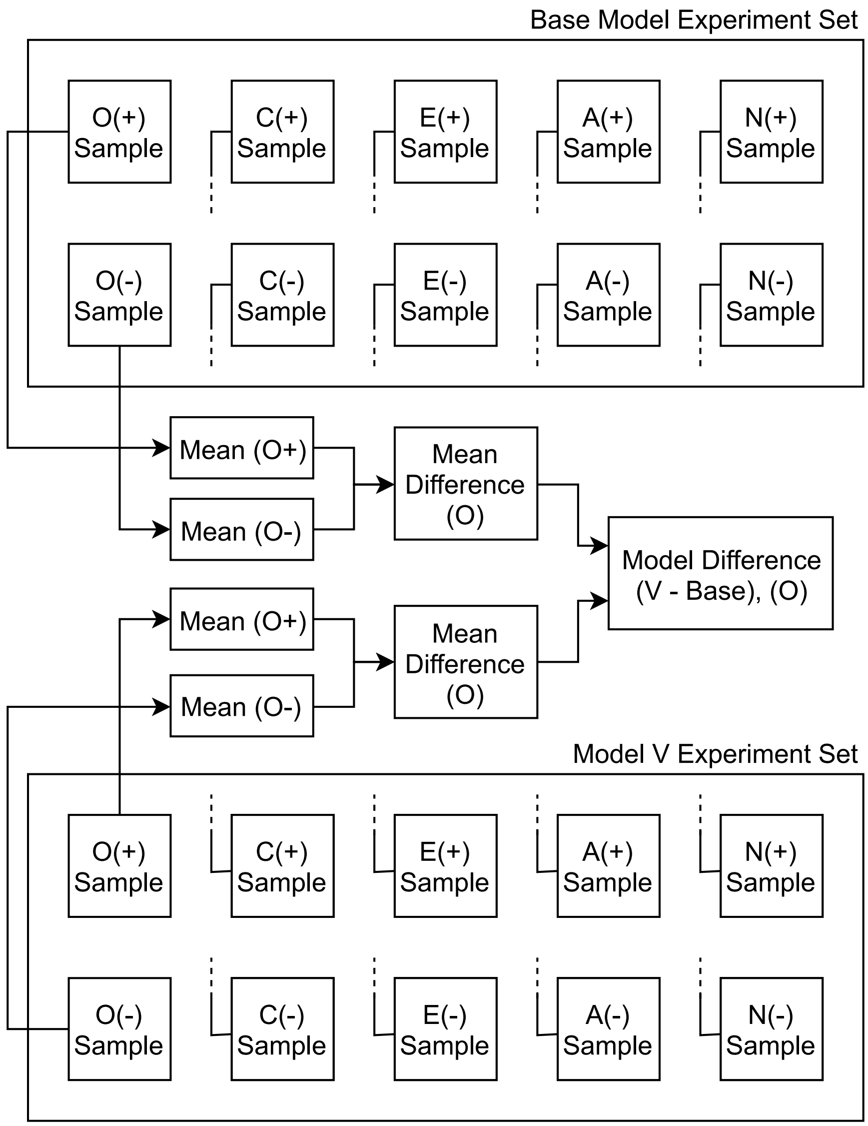


Figure 7.1: The comparisons used in Main Experiments.

7.1 Preliminary Experiments

These experiments are done to determine system parameters. Agent Neutrality Experiment aims to select an agent model with minimum personality bias to use in the experiments. Facial Expression - Personality Experiment aims to find a correlation between the facial expression of the agent and its perceived personality. We use the found relation to enhance the agent's personality with facial expressions. The participant and sample count of Preliminary Experiments are different from Main Experiments, but the same metrics are used.

7.1.1 Agent Neutrality Experiment

Using the same idle pose with no modifications, we expect a slightly different OCEAN score for each 3D human model, due to differences in appearance (see Figure 7.2). In the most neutral case, the agent should have a result of 0.5 for each OCEAN factor. Because of bias in personality, some 3D models could have different neutral OCEAN scores, which could influence the efficacy of the system. To select an agent model with minimal personality bias, we generated ten different 3D human models using Adobe Fuse [45] and took a screenshot of the upper body in the idle pose. Each sample is rated by 25 different users in this experiment.

We define the neutrality of an agent as $N_{OCEAN} = 1 - (\sum |OCEAN_x - 0.5|)/5$, $N_{OCEAN} = 1$ being the agent with all neutral OCEAN factors. We choose the agent in Figure 7.2 (b) as the most neutral agent, with a mean of $N_{OCEAN} = 0.966$, and the agent in Figure 7.2 (e) as the least neutral agent, with a mean of $N_{OCEAN} = 0.916$. Welch's t-test gives a p-value of 0.013 for neutrality differences of the agents in Figure 7.2 (b) and (e). Table 7.5 provides the significances of the neutrality differences between each agent pair using Welch's t-test. Figure 7.2 provides all agent images and their corresponding mean neutrality values used in this experiment. The agent in Figure 7.2 (b) is used in rest of the experiments.

(a) $N = 0.931$ (b) $N = 0.966$ (c) $N = 0.918$ (d) $N = 0.931$ (e) $N = 0.916$

(f) $N = 0.919$ (g) $N = 0.950$ (h) $N = 0.938$ (i) $N = 0.938$ (j) $N = 0.949$

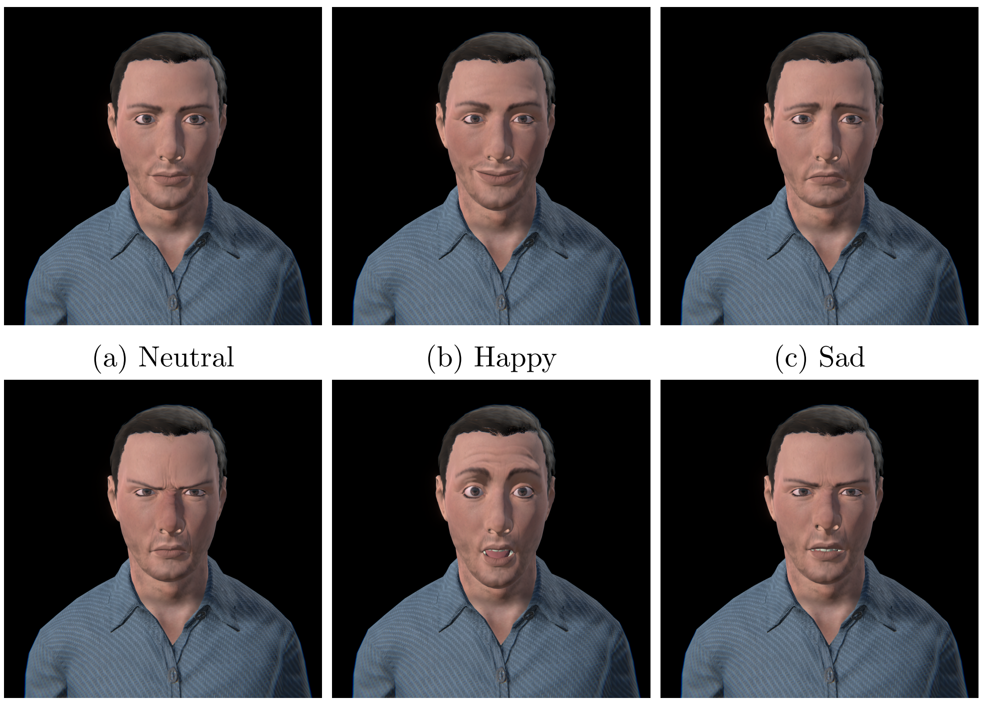
Figure 7.2: The 3D human models used in Agent Neutrality Experiment and the corresponding neutrality means.

Table 7.5: The significances of neutrality differences between agent pairs in Figure 7.2 using Welch's t-test. We do not expect a significant difference between all pairs; the existence of significant differences ($p < 0.05$) show that physical appearance by itself could influence the perceived personality.

Agent	**(a)**	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**	**(h)**	**(i)**	**(j)**
(a)	1.000	0.069	0.608	0.179	0.639	0.282	0.947	0.159	0.778	0.315
(b)	0.069	1.000	0.173	0.560	**0.013**	0.502	0.077	0.663	0.080	0.386
(c)	0.608	0.173	1.000	0.394	0.298	0.543	0.580	0.350	0.785	0.611
(d)	0.179	0.560	0.394	1.000	**0.048**	0.867	0.184	0.901	0.227	0.743
(e)	0.639	**0.013**	0.298	**0.048**	1.000	0.105	0.710	**0.043**	0.409	0.114
(f)	0.282	0.502	0.543	0.867	0.105	1.000	0.278	0.787	0.365	0.895
(g)	0.947	0.077	0.580	0.184	0.710	0.278	1.000	0.164	0.736	0.311
(h)	0.159	0.663	0.350	0.901	**0.043**	0.787	0.164	1.000	0.200	0.667
(i)	0.778	0.080	0.785	0.227	0.409	0.365	0.736	0.200	1.000	0.412
(j)	0.315	0.386	0.611	0.743	0.114	0.895	0.311	0.667	0.412	1.000

7.1.2 Facial Expression - Personality Experiment

We hypothesize certain facial expressions would influence the agent's perceived personality. To find out which facial expressions users relate to different personality factors, we made a user study based on images of the agent performing different facial expressions. The agent's facial expression is set to neutral, happy, sad, angry, surprised and disgusted in each sample. We show the agent's face as a close-up portrait. The images used in this experiment are given in Figure 7.3.

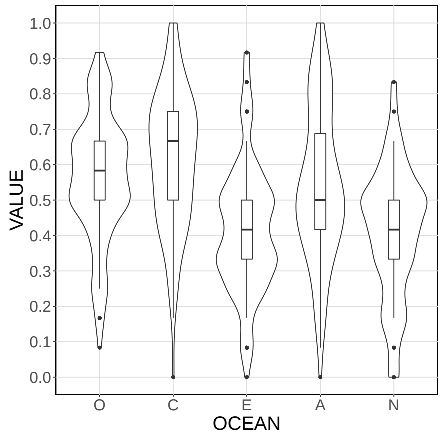


(a) Neutral (b) Happy (c) Sad

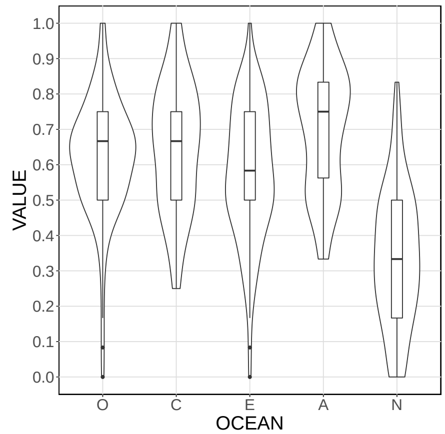
(d) Angry (e) Surprised (f) Disgusted

Figure 7.3: The facial expressions used in Facial Expression - Personality Experiment.

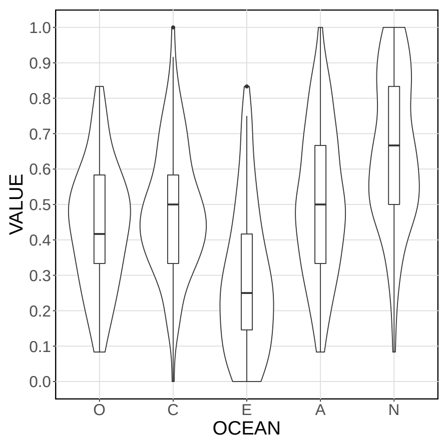
Each sample is rated by 100 different users in this experiment. Figure 7.4 depicts the OCEAN score distribution graph of each image. The results of the neutral expression show minor bias in openness, extraversion and neuroticism, and significant bias in conscientiousness. We take neutral expression's OCEAN score as a base value and compare other expressions to neutral expression to compensate this bias.



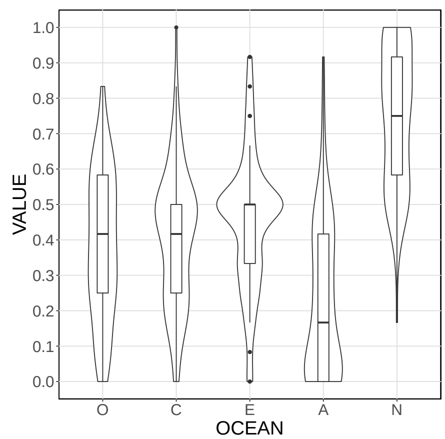
(a) Neutral



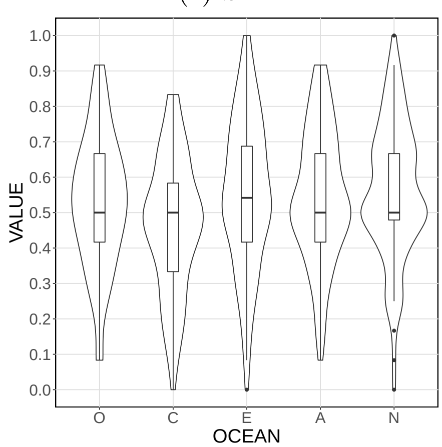
(b) Happy



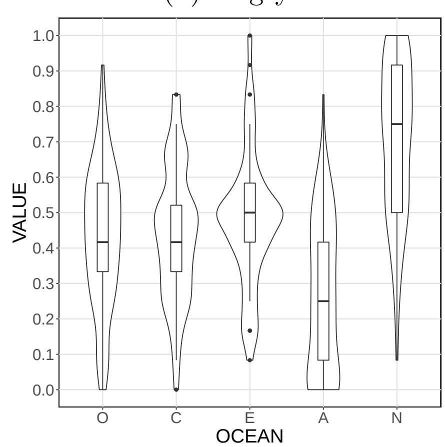
(c) Sad



(d) Angry



(e) Surprised



(f) Disgusted

Figure 7.4: The OCEAN score distribution graphs of samples from Facial Expression - Personality Experiment.

Table 7.6 show the mean differences of each expression to the neutral expression for each OCEAN factor. Some of the significant findings from this experiment are as follows. Happiness is most related to positive extraversion and positive agreeableness. Sadness is most related to positive neuroticism and negative extraversion. Anger is most related to negative agreeableness and positive neuroticism. Surprise is most related to negative conscientiousness and positive neuroticism. Disgust is most related to negative agreeableness and positive neuroticism.

Table 7.6: The mean differences of facial expressions to neutral expression. Statistically significant values ($p < 0.05$) are shown in bold.

Factor	**Happy**	**Sad**	**Angry**	**Surprised**	**Disgusted**
O	**0.057**	**-0.121**	**-0.177**	-0.089	**-0.128**
C	0.034	**-0.128**	**-0.203**	**-0.180**	**-0.175**
E	**0.157**	**-0.130**	0.037	**-0.020**	**0.063**
A	**0.155**	-0.033	**-0.311**	-0.152	**-0.279**
N	**-0.074**	**0.236**	**0.330**	**0.222**	**0.289**

Table 7.7 provides the significance of the difference of each facial expression compared to the neutral expression as p-values for Welch's t-test. The results support this experiment's hypothesis about facial expressions being influential on the perceived personality. Based on the significant differences, we define additive facial expression values for each OCEAN factor extreme, given in Table 7.8. When an agent with certain personality speaks a dialogue unit, it shows an expression calculated using Watson NLP. Additive facial expression values are added to the calculated expression based on the agent's OCEAN personality.

Table 7.7: The significance of each expression's difference from neutral expression, using Welch's t-test. Statistically significant values ($p < 0.05$) are shown in bold.

Factor	**Happy**	**Sad**	**Angry**	**Surprised**	**Disgusted**
O	**0.019**	**<0.001**	**<0.001**	0.223	**<0.001**
C	0.199	**<0.001**	**<0.001**	**<0.001**	**<0.001**
E	**<0.001**	**<0.001**	0.166	**<0.001**	**0.016**
A	**<0.001**	0.267	**<0.001**	0.907	**<0.001**
N	**0.005**	**<0.001**	**<0.001**	<0.001	**<0.001**

Using the additive expression values, a positive agreeableness agent would increase happy expression, and reduce angry, surprised and disgusted expressions of each dialogue unit. This agent would still perform angry, surprised or disgusted expressions, however, the expression value that is calculated using Watson NLP should be high enough to surpass the negative additive values. Each facial expression could take a value between 0 and 1, and the resulting expression of the agent is a blend of all non-zero expressions.

Negative additive expression values could only decrease the related expression to 0. We clamp the expression values between 0 and 1 for each dialogue unit. Increasing happy expression does not decrease sadness directly, we, therefore, define a negative additive value in certain cases. The values in Table 7.8 are related to the mean differences given in Table 7.6, however, we adjust values to avoid unnatural blends in certain cases.

Table 7.8: Additive facial expression values, decided using the significant differences in Facial Expression - Personality Experiment. Each cell contains the additive adjustment value to the expression calculated using Watson NLP for a dialogue unit.

Sample	**Happy**	**Sad**	**Angry**	**Surprise**	**Disgust**
O (+)	0.125	0.000	-0.250	0.000	0.000
O (−)	-0.125	0.000	0.250	0.000	0.000
C (+)	0.125	0.000	-0.125	-0.350	-0.125
C (−)	-0.125	0.000	0.125	0.350	0.125
E (+)	0.250	-0.250	0.000	0.000	0.000
E (−)	-0.250	0.250	0.000	0.000	0.000
A (+)	0.250	0.000	-0.500	-0.250	-0.350
A (−)	-0.250	0.000	0.500	0.250	0.350
N (+)	-0.125	0.350	0.500	0.350	0.350
N (−)	0.125	-0.350	-0.500	-0.350	-0.350

In this experiment, we do not test for the compound expressions, such as the combination of surprised and happy; however, in the main experiments when additive expressions are used, these blends occur. A surprised and happy expression combination could relate to different personality factors that are not included in the sum of surprised and happy, therefore testing compound expressions remains a future research interest. We did not include afraid expression because it was not distinguishable from surprise expression using the facial shape keys of the current human model.

7.2 Main Experiments

In each Main Experiment, we test an implemented model having different modalities. One experiment includes 10 samples using the model of interest, rated by 50 different users. For each experiment, we choose non-overlapping participants to prevent the bias caused by the user getting familiar with the agents. We use the Passport Scenario without changing the flow of a sample's dialogue between different models, however, the answers of agents are determined by the OCEAN factors. According to the sample, one OCEAN factor of interest is set to the positive or negative extreme while remaining factors stay neutral. Agent names and occupations are randomized to suggest that each sample contains a different agent. The same agent model is used for all visual experiments. When there is a visible agent model, Oculus Lipsync is used to match the agent mouth motion to the spoken sound.

The results table given for each experiment contain one OCEAN factor of interest in each row. The means of positive and negative samples are compared for that OCEAN factor, and the mean difference is given as well as the p-value of a Welch's T-Test. We test for the null hypothesis and p-values less than 0.05 indicate a significant difference between the results of positive and negative samples. The mean difference is also used as a metric to compare the results of different experiments. The higher the mean difference is, the better positive and negative samples are distinguished.

We compared the results of different models using a one-way ANOVA test at the end to reveal significant differences per OCEAN factor. This test is based on positive and negative sample differences. We expect the mean differences to be higher as more modules are active. Models are designed as shown in Figure 7.5. In each model, the agent logic that handles the scenario is active, as well as the dialogue module that is responsible for choosing dialogue text based on personality. Base Model does not have any other module. Model V makes appropriate voice changes based on the OCEAN personality of the agent, following [4]. Model VF applies facial expression changes, using both Watson NLP API and additive expressions based on the agent personality. Each time the agent starts talking, an emotion is calculated using Watson NLP and mapped into a facial expression. Based on additive facial expression values, introduced in Table 7.8, the facial expression is further adjusted to fit the agent's personality. While the agent is listening, its facial expression decays with time.

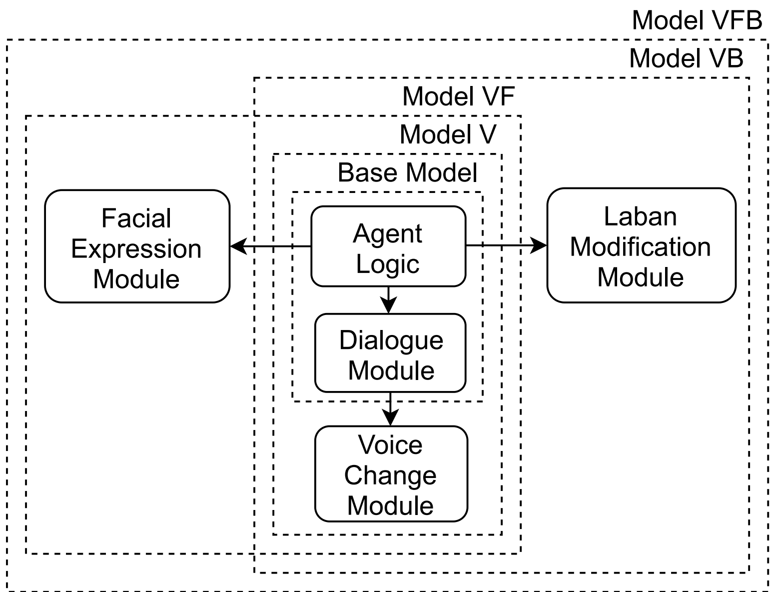


Figure 7.5: Models used in main experiments and modules included in each model.

7.2.1 Base Model

In this experiment, we test the influence of the dialogue text, that is crafted based on Mairesse features [5]. The voice of the speaking agent remains the same for all samples. The role of the speaker, as well as the spoken text, is visible on the screen. There are no other graphical elements on the black background, as seen in Figure 7.6. In each sample, the dialogue of the visitor (the agent of interest) and the passport officer (the user representative) is vocalized using the Watson Text to Speech service.

Table 7.9 show that the dialogue content by itself could create significant differences in perceived personality. Agreeableness is distinguished the best, probably due to the frequent usage of respect and hate words. Extraversion is distinguished

(a) Positive Extraversion Sample

(b) Negative Extraversion Sample

Figure 7.6: The visuals from Base Model.

the least; this could be because of the formality the scenario requires. Friendliness of the positive extraversion sample and the quietness of the negative extraversion sample is not emphasized enough in the dialogue text.

This model is used as a base case and we aim to achieve better mean differences by adding vocal adjustments and graphical cues in the following experiments.

Table 7.9: Welch's t-test results for null hypothesis between positive and negative samples per OCEAN factor using Base Model. Bold values indicate significant differences ($p < 0.05$).

Ocean Factor	Mean (+)	Mean (-)	Mean Difference	Difference P-Value
O	0.610	0.531	0.079	**0.003**
C	0.560	0.496	0.064	**0.045**
E	0.543	0.516	0.027	0.378
A	0.561	0.458	0.103	**0.001**
N	0.540	0.443	0.097	**0.004**

7.2.2 Model V: Voice Adjustment

In this experiment, we use the voice-altering system of the Watson Text to Speech service [43], following the OCEAN personality - vocal features mapping of [4]. The dialogue content is shown on the screen, the same as Base Model shown in Figure 7.6. The voice of the passport officer is kept the same.

Table 7.10: Welch's t-test results for null hypothesis between positive and negative samples per OCEAN factor using Model V. Bold values indicate significant differences ($p < 0.05$).

Ocean Factor	**Mean (+)**	**Mean (-)**	**Mean Difference**	**Difference P-Value**
O	0.610	0.518	0.092	**0.023**
C	0.596	0.46	0.136	**0.001**
E	0.578	0.421	0.157	**<0.001**
A	0.621	0.436	0.185	**<0.001**
N	0.570	0.395	0.175	**<0.001**

As shown in Tables 7.10 and 7.11, altering the speech slightly improves the mean difference. Extraversion is improved the most, while Openness is the least influenced factor. Neuroticism is the most differentiated factor in this experiment. Table 7.11 shows the mean differences of each model per OCEAN factor. Because Table 7.15 on page 68 shows the pairwise comparison of all models includes ANOVA results for every model, we do not apply a separate pairwise test in this part. Pairwise test results do not show a significant difference between Base Model and Model V, however, Model V distinguished every factor better than Base Model. The voice-altering system is kept active in the remaining experiments.

7.2.3 Model VF: Voice and Facial Expression Adjustment

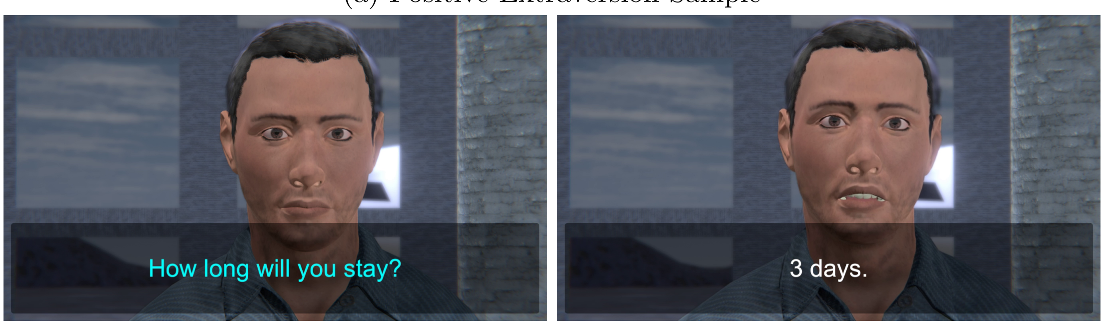
In this experiment, on top of the vocal adjustments, we activate the facial expression module. In each sample, during the agent's line of dialogue, the facial expression of the agent is updated using Watson NLP and additive expression

Table 7.11: The mean differences of positive and negative samples for each model per OCEAN factor. Higher values correspond to better distinguished factors.

Model	**O**	**C**	**E**	**A**	**N**
Base	0.079	0.064	0.027	0.103	0.097
V	0.092	0.136	0.157	0.185	0.175
VF	0.142	0.196	0.260	0.302	0.340
VB	0.193	0.290	0.353	0.387	0.357
VFB	0.255	0.250	0.494	0.442	0.460

values defined in Facial Expression - Personality Experiment. The camera is focused on the agent's face, as shown in Figure 7.7. The animation of the agent is unaltered, only the facial expression changes and the eyes blink.

(a) Positive Extraversion Sample



(b) Negative Extraversion Sample

Figure 7.7: The visuals from VF Model.

Tables 7.11 and 7.12 show that the mean difference is improved the most for extraversion, agreeableness, and neuroticism. This could be explained by facial expressions being more influential on these three OCEAN factors. Table 7.15 reveals that the results of Model VF are significantly better than those of Base Model for extraversion and neuroticism. The worst performing factor for

Model VF is openness, followed by conscientiousness.

Table 7.12: Welch's t-test results for null hypothesis between positive and negative samples per OCEAN factor using Model VF. Bold values indicate significant differences ($p < 0.05$).

Ocean Factor	**Mean (+)**	**Mean (-)**	**Mean Difference**	**Difference P-Value**
O	0.613	0.471	0.142	**0.001**
C	0.621	0.425	0.196	**<0.001**
E	0.635	0.375	0.260	**<0.001**
A	0.643	0.341	0.302	**<0.001**
N	0.641	0.301	0.340	**<0.001**

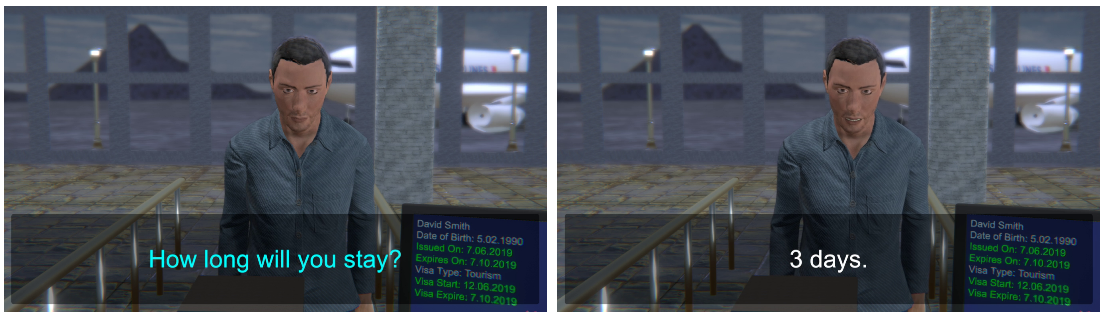
We consider users do not make strong judgments based on gestural cues for openness and conscientiousness. Users may need to see the agent for longer durations in different situations because these two factors are more intelligence-related and situational. It is hard to show the openness of an agent in a limited scenario. An agent with high openness would try new options in similar circumstances, while an agent with low openness would stick with the same option. A strong judgment on conscientiousness may require a trust-based relation with the user, and relying purely on visual cues could be confusing because the uncareful nature of low conscientiousness could be confused with high energy, and deliberation of high conscientiousness could be confused with low energy traits.

7.2.4 Model VB: Voice and Body Adjustment

In this experiment, we use movement modifications based on LMA to alter agent animations to express a certain personality. The altered voice for the agent is used with no facial expressions, only the eyes of the agent blinks. The upper body of the agent is visible on the screen, for users to be able to focus on the agent's movements, as seen in Figure 7.8.

Tables 7.11 and 7.13 show that the results are improved the best for extraversion, agreeableness, and conscientiousness as compared to Model VF. Although the significance of their difference is low, this experiment shows adjusting body

(a) Positive Extraversion Sample



(b) Negative Extraversion Sample

Figure 7.8: The visuals from VB Model.

movements is a better candidate for expressing personality than to adjust facial expressions when applicable. However, certain scenes would not allow showing the whole body of an agent. In such cases, Model VF would be the better candidate. Because Model VF and Model VB have non-overlapping active modules, they can be combined into a better model, discussed in the next experiment.

Model VB is significantly better than Base Model in terms of every OCEAN factor, except for openness. This model best distinguishes extraversion, agreeableness, and neuroticism; while the worst-performing factor is openness followed by conscientiousness, which supports the findings of [32] that uses LMA adjustments to express personality.

Table 7.13: Welch's t-test results for null hypothesis between positive and negative samples per OCEAN factor using Model VB. Bold values indicate significant differences ($p < 0.05$).

Ocean Factor	**Mean (+)**	**Mean (-)**	**Mean Difference**	**Difference P-Value**
O	0.668	0.475	0.193	**<0.001**
C	0.698	0.408	0.290	**<0.001**
E	0.693	0.340	0.353	**<0.001**
A	0.695	0.308	0.387	**<0.001**
N	0.655	0.298	0.357	**<0.001**

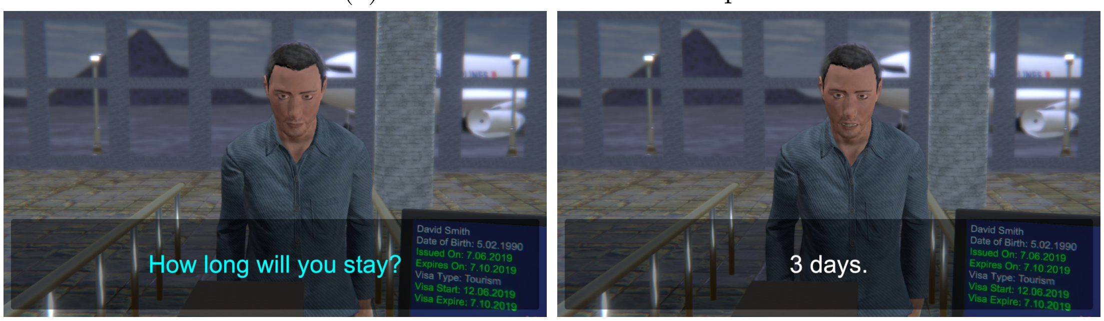
7.2.5 Model VFB: Voice, Facial Expression, and Body Adjustment

In this experiment, all implemented modules are active, including voice, movement, and facial expression adjustment modules. The upper half of the agent's body is visible, similar to Model VB, to be able to show limb movements, as seen in Figure 7.9. On the other hand, since the agent's face covers a smaller portion of the screen compared to Model VF, improvement caused by the facial expression module is not expected to be at its full potential.

Tables 7.11 and 7.14 show that overall this is the best performing model. Table 7.15 reveals Model VFB is significantly better than Base Model for all factors, and better than Model V for all factors, except for conscientiousness. Furthermore, Model VFB is significantly better than Model VF in terms of extraversion. Using facial expressions with movement adjustments improves the mean differences of every factor except for conscientiousness compared to Model VB. There is a slight decrease in the conscientiousness of Model VFB compared to Model VB, that is not significant ($p = 0.954$). We consider this to be related to conscientiousness being the least influenced factor by facial expressions.

Table 7.16 shows the differences between the mean differences of pairs of models. Each model is compared to a preceding model and positive values correspond to improvements. Only one step contains a slight decrease, which is not significant. Figures 7.10 (a) and (b) show the mean differences of positive and negative

(a) Positive Extraversion Sample



(b) Negative Extraversion Sample

Figure 7.9: Visuals from Model VFB.

Table 7.14: Welch's t-test results for null hypothesis between positive and negative samples per OCEAN factor using Model VFB. Bold values indicate significant differences ($p < 0.05$).

Ocean Factor	**Mean (+)**	**Mean (-)**	**Mean Difference**	**Difference P-Value**
O	0.738	0.483	0.255	**<0.001**
C	0.713	0.463	0.250	**<0.001**
E	0.785	0.291	0.494	**<0.001**
A	0.723	0.281	0.442	**<0.001**
N	0.685	0.225	0.460	**<0.001**

Table 7.15: The results of one-way ANOVA test with Tukey HSD for Mean Differences of different models for each OCEAN factor. Bold values show significant differences ($p < 0.05$).

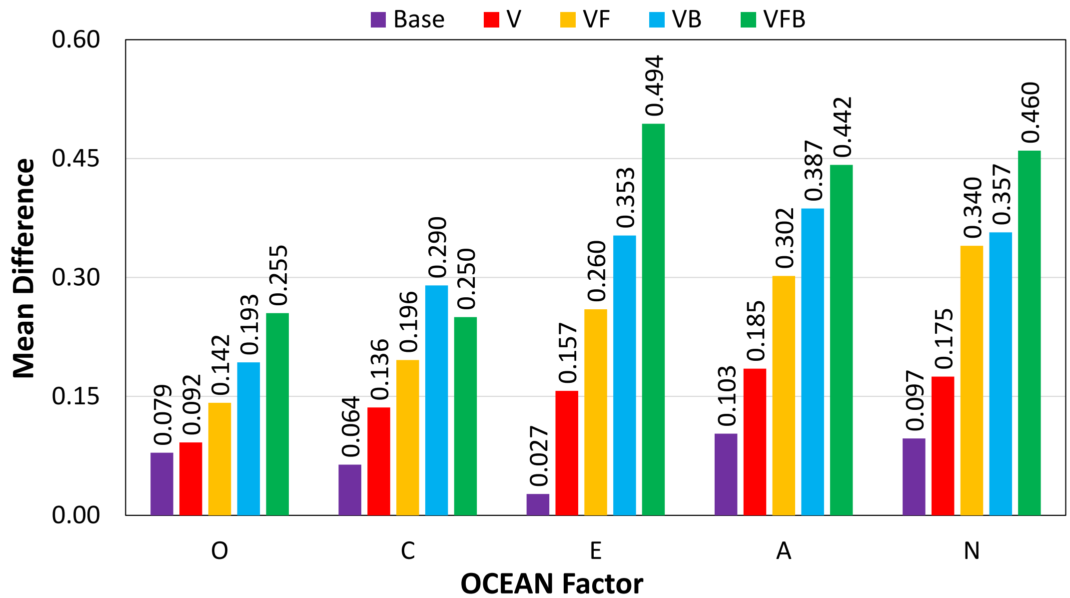
Model	**O**	**C**	**E**	**A**	**N**
V - Base	0.999	0.690	0.293	0.813	0.763
VF - Base	0.756	0.128	**0.005**	0.066	**0.002**
VB - Base	0.197	**<0.001**	**<0.001**	**0.001**	**0.001**
VFB - Base	**0.009**	**0.009**	**<0.001**	**<0.001**	**<0.001**
VF - V	0.881	0.824	0.530	0.530	0.097
VB - V	0.314	0.053	**0.028**	0.059	0.051
VFB - V	**0.019**	0.263	**<0.001**	**0.006**	**<0.001**
VB - VF	0.867	0.463	0.627	0.790	0.999
VFB - VF	0.210	0.878	**0.004**	0.341	0.371
VFB - VB	0.774	0.954	0.222	0.949	0.527

samples and their significances in each model for each OCEAN factor, respectively. Model VFB is shown to be the best performing one in distinguishing opposite samples for every factor, except conscientiousness, which is because of conscientiousness is not reflected well to facial expressions.

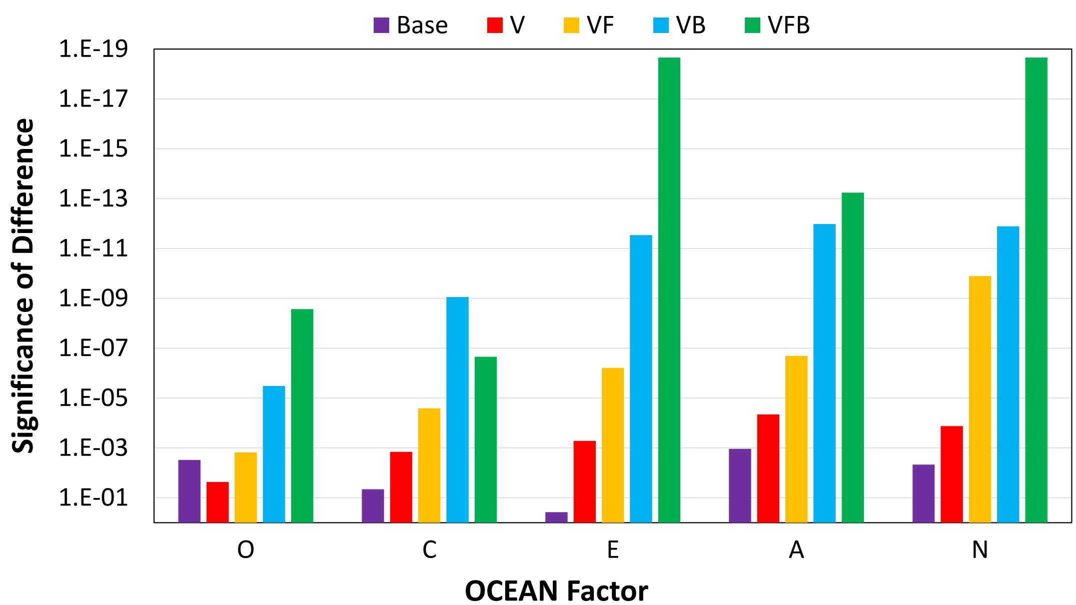
Table 7.16: The mean differences of different models for each OCEAN factor. Bold values show significant differences ($p < 0.05$).

Model	**O**	**C**	**E**	**A**	**N**
V - Base	0.013	0.073	0.130	0.081	0.078
VF - Base	0.063	0.133	**0.233**	0.198	**0.243**
VB - Base	0.115	**0.226**	**0.326**	**0.283**	**0.260**
VFB - Base	**0.176**	**0.186**	**0.466**	**0.338**	**0.363**
VF - V	0.05	0.060	0.103	0.116	0.165
VB - V	0.101	0.153	**0.196**	0.201	0.181
VFB - V	**0.163**	0.113	**0.336**	**0.256**	**0.285**
VB - VF	0.051	0.093	0.093	0.085	0.016
VFB - VF	0.113	0.053	**0.233**	0.140	0.120
VFB - VB	0.061	-0.040	0.140	0.055	0.103

Combining facial expressions with movement modifications works best for extraversion, this is probably due to each module adding non-overlapping features related to extraversion. With facial expressions, positive emotions and friendliness are projected. With movement modifications, energetic and assertive nature



(a)



(b)

Figure 7.10: The comparison of different models for each OCEAN factor: (a) mean differences of positive and negative samples (higher is better) and (b) significances (lower is better) for each factor. The graph in (b) is in log scale.

of positive extraversion is shown. For agreeableness, both modules express similar features of sympathy and modesty, therefore their combined version is not as effective as extraversion.

To show the mood changing nature of positive neuroticism additive facial expressions could be multiplied with randomized values so that a similar dialogue unit would have different facial expressions each time. However, such a mechanism would require longer exposure to the user to show the influence on perceived personality. Features such as dependability of conscientiousness and adventurous nature of high openness also require a longer exposure of the agent, and such traits could be expressed through dialogue flow. This would require behavioral planning such as in [35], that would influence agent decisions based on personality.

Because all OCEAN scores are normalized, the maximum possible difference for positive and negative samples would be 1 theoretically; however, how strong users could make personality judgments based on a short video clip remains a future research topic. It might not be possible to distinguish agent personality higher than a threshold. Comparing the behavior of the agent to that of a person in real videos of similar scenarios could also be used as an evaluation metric, however, this would require a suitable dataset with personality labels.

Chapter 8

Conclusion and Future Research Possibilities

We propose a framework to generate a conversational agent that expresses personality through vocal and visual cues. Animation modification, facial expression adjustment, voice transformation, and dialogue selection are used in cooperation to express certain personality types. We perform user studies to evaluate the different modules used in the architecture and show that the combination of the implemented modules performs the best overall. We use the distinction between positive and negative samples per OCEAN factor as an evaluation metric.

The results show that using the implemented system extraversion is distinguished the best, followed by neuroticism and agreeableness, due to the expressive nature of these factors. Openness and conscientiousness are not distinguished as much, because the cues of these factors are more intelligence-based and require exposure to the agent in different situations.

Based on the results, we could conclude that each introduced model could be used to distinguish between opposite personality types; however, the combination of facial expression adjustment module with animation modification and vocal transformation modules perform significantly better. This suggests combining

LMA modifications with facial expressions is an effective way of expressing agent personality, and should be adopted where it is applicable. In some cases, the scenes where the agent's face or body is not visible could decrease the efficacy of the solution, and in some others, the agent could be unable to express facial expressions. In all other cases, we consider Model VFB to be more effective.

We performed user studies using recordings of the agent, however, we did not experiment with the case where the user directly interacts with the agent. We aim to test the system with users directly interacting with the agent through conversation. Direct interaction could reveal different results because the user would feel more empathy for the agent. It would also be interesting to see how the user chooses his or her dialogue sentences based on the agent's personality. The user might talk more politely to a high agreeableness agent and might not be interested in a conversation with a low extraversion agent.

Some future research possibilities are as follows.

Personality-based automated style transfer for text: Currently, we use handcrafted OCEAN Alternatives for dialogue; however, it could be possible to inject specific personality into a neutral sentence. This operation would require large data sets with sentences that have a similar meaning and different personality scores for training. Several text corpora with matching personality scores exist, yet they do not focus on single sentences. Existing systems for automated personality analysis from text, such as Watson Natural Language Understanding API [43], require longer paragraphs with many sentences. Having multiple units that focus on singular text operations, based on Mairesse features could be a possible solution. For example, one unit could be responsible for adding hesitation into text by a given degree, and personality could be mapped into such units.

Machine learning for animation modification: Given motion capture data with personality scores, mapping between current animation parameters and personality can be found in a better way. Furthermore, with a large dataset containing animations of similar tasks by different people with varying personality scores, it could be possible to leave the modification process to a neural network. However, constructing such dataset would be a hard task, and differences could be very subtle that would require a high-end motion capture setup. Facial expressions could also be captured to accompany movement modifications.

A sophisticated model for agent emotions: The agent emotions are currently reflected as facial expressions that are adjusted based on the agent's dialogue sentences and personality. Agent emotions could include the influence of user speech, so that toning and word selection of the user would have an impact on the agent's facial expression. For example, if the user speaks in an angry manner, the agent could act more defensively, and in case the user speaks friendly, the agent could express in a similar way. The implementation of such a system would require modeling of emotions for short speech clips, as well as a conversational emotion mapping between speakers. This emotion mapping would potentially be influenced by personality as well. For example, a high agreeableness agent could be more caring when the user sounds sad, as opposed to a low agreeableness agent that might not be interested in cheering the user. The facial expression of the user could also be captured to classify user emotion.

Interaction with multiple agents: Currently, the system focuses on conversation with one agent. It could be interesting to integrate it into a simulation where multiple agents interact based on personality. The goals and needs for each agent could be determined by a model, and human players in such simulation could train agent behavior. In a multi-user scenario, users would be given a task to perform, and the avatar of the user would have needs to fulfill. User behavior in the simulated world would be recorded, and a personality assessment test to the user would be made at the end. A mapping between personality and behavior could be found for using it with simulated agents.

An animation modification process for 2D: It could be interesting to design an alternative animation modification process for 2D figures. Reduced dimensionality would require changes in the application of LMA. Due to limitations, exaggerating the movement might be necessary. Facial expressions on a cartoon-like character could have different influences. To our knowledge, there are no solutions that focus on 2D figures and the limitations in this area require further exploration.

Nonverbal interaction: Currently, the user interacts with the agent through speech only. Nonverbal interaction could be introduced by tracking user motion. The user could point or use gesture to communicate with the agent in a different way.

The modular structure of the implemented system makes it possible to easily replace individual components such as movement modifiers, personality mappers, and facial expression adjustment module; thus, the existing architecture could be used in further research.

Bibliography

[1] Paul Ekman International, "The 5 Communication Channels." https://www.ekmaninternational.com/wiki/the-5-communication-channels/. Accessed: 2019-05-14.

[2] A. Todorov, C. P. Said, A. D. Engell, and N. N. Oosterhof, "Understanding Evaluation of Faces on Social Dimensions," *Trends in Cognitive Sciences*, vol. 12, no. 12, pp. 455–460, 2008.

[3] L. P. Naumann, S. Vazire, P. J. Rentfrow, and S. D. Gosling, "Personality Judgments Based on Physical Appearance," *Personality and Social Psychology Bulletin*, vol. 35, no. 12, pp. 1661–1671, 2009.

[4] T. Polzehl, *Personality in Speech: Assessment and Automatic Classification*. Cham, Switzerland: Springer International Publishing, 2015.

[5] F. Mairesse, *Learning to Adapt in Dialogue Systems: Data-driven Models for Personality Recognition and Generation*. PhD thesis, Department of Computer Science, University of Sheffield, Sheffield, UK, February 2008.

[6] M. Neff, Y. Wang, R. Abbott, and M. Walker, "Evaluating the Effect of Gesture and Language on Personality Perception in Conversational Agents," in *Proceedings of the International Conference on Intelligent Virtual Agents*, IVA '10, (Berlin, Germany), pp. 222–235, Springer, 2010.

[7] F. Durupinar, M. Kapadia, S. Deutsch, M. Neff, and N. I. Badler, "PERFORM: Perceptual Approach for Adding OCEAN Personality to Human Motion Using Laban Movement Analysis," *ACM Transactions on Graphics*, vol. 36, no. 1, p. 6, 2017.

[8] R. R. McCrae and O. P. John, "An Introduction to the Five-factor Model and Its Applications," *Journal of Personality*, vol. 60, no. 2, pp. 175–215, 1992.

[9] S. D. Gosling, "Personality in Non-human Animals," *Social and Personality Psychology Compass*, vol. 2, no. 2, pp. 985–1001, 2008.

[10] R. D. Latzman, W. D. Hopkins, A. C. Keebaugh, and L. J. Young, "Personality in Chimpanzees (Pan Troglodytes): Exploring the Hierarchical Structure and Associations with the Vasopressin V1A Receptor Gene," *PLoS One*, vol. 9, no. 4, Article no. e95741, 2014.

[11] V. Maletic, *Body-Space-Expression: The Development of Rudolf Laban's Movement and Dance Concepts*, vol. 75. Berlin, Germany: Walter de Gruyter, 2011.

[12] B. Adrian, *Actor Training the Laban Way: An Integrated Approach to Voice, Speech, and Movement.* New York, NY, USA: Allworth Press, 2008.

[13] S. J. Burton, A.-A. Samadani, R. Gorbet, and D. Kulić, "Laban Movement Analysis and Affective Movement Generation for Robots and Other Near-living Creatures," in *Dance Notations and Robot Motion*, pp. 25–48, Cham, Switzerland: Springer International Publishing, 2016.

[14] R. Bernstein, T. Shafir, R. Tsachor, K. Studd, and A. Schuster, "Laban Movement Analysis Using Kinect," *International Journal of Computer and Information Engineering*, vol. 9, pp. 1394–1398, 2015.

[15] T. Lourens, R. Van Berkel, and E. Barakova, "Communicating Emotions and Mental States to Robots in a Real time Parallel Framework Using Laban Movement Analysis," *Robotics and Autonomous Systems*, vol. 58, no. 12, pp. 1256–1265, 2010.

[16] L. Zhao and N. I. Badler, "Synthesis and Acquisition of Laban Movement Analysis Qualitative Parameters for Communicative Gestures," Tech. Rep. IRCS-01-11, The Institute for Research in Cognitive Science (IRCS), University of Pennsylvania, 2001.

[17] C. Larboulette and S. Gibet, “A Review of Computable Expressive Descriptors of Human Motion,” in *Proceedings of the 2nd International Workshop on Movement and Computing*, MOCO ’15, (New York, NY, USA), pp. 21–28, ACM, 2015.

[18] D. E. Matsumoto, H. C. Hwang, and M. G. Frank, *APA Handbook of Nonverbal Communication*. Washington, DC, USA: American Psychological Association, 2016.

[19] M. L. Knapp, J. A. Hall, and T. G. Horgan, *Nonverbal Communication in Human Interaction*. Boston, MA, USA: Cengage Learning, 2013.

[20] F. Poyatos, *Nonverbal Communication Across Disciplines: Volume 2: Paralanguage, Kinesics, Silence, Personal and Environmental Interaction*. Amsterdam, The Netherlands: John Benjamins Publishing, 2002.

[21] D. C. Funder, *The Personality Puzzle: Seventh International Student Edition*. New York, NY, USA: WW Norton & Company, 2015.

[22] J. Allbeck and N. Badler, “Toward Representing Agent Behaviors Modified by Personality and Emotion,” in *Proceedings of the Workshop on Embodied Conversational Agents at the First International Conference on Autonomous Agents and Multiagent Systems*, vol. 2 of *AAMAS ’02*, pp. 15–19, 2002.

[23] D. Chi, M. Costa, L. Zhao, and N. Badler, “The EMOTE Model for Effort and Shape,” in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH ’00, (New York, NY, USA), pp. 173–182, ACM, 2000.

[24] J. A. Levy and M. P. Duke, “The Use of Laban Movement Analysis in the Study of Personality, Emotional State and Movement Style: An Exploratory Investigation of the Veridicality of “Body Language”,” *Individual Differences Research*, vol. 1, no. 1, pp. 39–63, 2003.

[25] F. Mairesse and M. A. Walker, “Towards Personality-based User Adaptation: Psychologically Informed Stylistic Language Generation,” *User Modeling and User-Adapted Interaction*, vol. 20, no. 3, pp. 227–278, 2010.

[26] Y. R. Tausczik and J. W. Pennebaker, "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology*, vol. 29, no. 1, pp. 24–54, 2010.

[27] M. R. Key, *Paralanguage and Kinesics (Nonverbal Communication)*. Lanham, MD, USA: The Scarecrow Press, Inc., 1975.

[28] R. R. McCrae and P. T. Costa, *Personality in Adulthood: A Five-factor Theory Perspective*. New York, NY, USA: Guilford Press, 2005.

[29] B. Ran, S. Tal, T. Rachelle, S. Karen, and S. Assaf, "Multitask Learning for Laban Movement Analysis," in *Proceedings of the 2nd International Workshop on Movement and Computing*, MOCO '15, (New York, NY, USA), pp. 37–44, ACM, 2015.

[30] A. Camurri, I. Lagerlöf, and G. Volpe, "Recognizing Emotion from Dance Movement: Comparison of Spectator Recognition and Automated Techniques," *International Journal of Human-Computer Studies*, vol. 59, no. 1-2, pp. 213–225, 2003.

[31] A. Aristidou, P. Charalambous, and Y. Chrysanthou, "Emotion Analysis and Classification: Understanding the Performers' Emotions Using the LMA Entities," *Computer Graphics Forum*, vol. 34, no. 6, pp. 262–276, 2015.

[32] H. J. Smith and M. Neff, "Understanding the Impact of Animated Gesture Performance on Personality Perceptions," *ACM Transactions on Graphics*, vol. 36, no. 4, p. 49, 2017.

[33] M. Masuda and S. Kato, "Motion Rendering System for Emotion Expression of Human Form Robots Based on Laban Movement Analysis," in *Proceedings of the 19th IEEE International Symposium on Robot andHuman Interactive Communication*, RO-MAN 10, pp. 324–329, IEEE, 2010.

[34] T. Randhavane, A. Bera, K. Kapsaskis, R. Sheth, K. Gray, and D. Manocha, "EVA: Generating Emotional Behavior of Virtual Agents using Expressive Features of Gait and Gaze," in *Proceedings of the ACM Symposium on Applied Perception*, SAP '19, (New York, NY, USA), ACM, 2019.

[35] M. Shvo, J. Buhmann, and M. Kapadia, "An Interdependent Model of Personality, Motivation, Emotion, and Mood for Intelligent Virtual Agents," in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, IVA '19, (New York, NY, USA), pp. 65–72, ACM, 2019.

[36] J. Golbeck, C. Robles, and K. Turner, "Predicting Personality with Social Media," in *ACM CHI Conference on Human Factors in Computing Systems, Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, (New York, NY, USA), pp. 253–262, ACM, 2011.

[37] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text," *Journal of Artificial Intelligence Research*, vol. 30, pp. 457–500, 2007.

[38] F. Mairesse and M. Walker, "PERSONAGE: Personality Generation for Dialogue," in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 496–503, 2007.

[39] P. T. Costa and R. R. McCrae, *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI): Professional Manual*. Odessa, FL, USA: Psychological Assessment Resources, Inc., 1992.

[40] L. H. Gilpin, D. M. Olson, and T. Alrashed, "Perception of Speaker Personality Traits Using Speech Signals," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '18, (New York, NY, USA), ACM, 2018.

[41] C. Charalambous, Z. Yumak, and A. F. van der Stappen, "Audio-driven Emotional Speech Animation for Interactive Virtual Characters," *Computer Animation and Virtual Worlds*, vol. 30, no. 3-4, Article no. e1892, 11 pages, 2019.

[42] Unity Technologies, "Unity." `https://unity.com/`. Accessed: 2019-05-05.

[43] International Business Machines Corporation, "IBM Watson API." `https://www.ibm.com/watson`. Accessed: 2019-05-04.

[44] Oculus VR, LLC, "Oculus Lipsync." `https://developer.oculus.com/downloads/package/oculus-lipsync-unity/`. Accessed: 2019-05-04.

[45] Adobe, Inc., "Adobe Fuse CC." `https://www.adobe.com/tr/products/fuse.html`. Accessed: 2019-05-05.

[46] S. Hoppe, T. Loetscher, S. A. Morey, and A. Bulling, "Eye movements during everyday behavior predict personality traits," *Frontiers in Human Neuroscience*, vol. 12, Article no. 105, 5 pages, 2018.

[47] Amazon.com, Inc., "Amazon Mechanical Turk." `https://www.mturk.com/`. Accessed: 2019-05-05.

[48] S. D. Gosling, P. J. Rentfrow, and W. B. Swann Jr, "A Very Brief Measure of the Big-Five Personality Domains," *Journal of Research in Personality*, vol. 37, no. 6, pp. 504–528, 2003.

[49] J. Tukey, "Comparing Individual Means in the Analysis of Variance," *Biometrics*, vol. 5, no. 2, pp. 99–114, 1949.

[50] A. Stoll, "Post Hoc Tests: Tukey Honestly Significant Difference Test," in *The SAGE Encyclopedia of Communication Research Methods* (M. Allen, ed.), pp. 1306–1307, Thousand Oaks, CA: SAGE Publications, Inc, 2018.