Department of Computing

An Affective Personality for an Embodied Conversational Agent

Xiao He

This thesis is presented for the Degree of
Master of Computer Science
of
Curtin University of Technology

November 2007

Declaration

This thesis contains no material that has been accepted for the award of any other degree or diploma in any university.

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

Signature: _____

Date: _____

Abstract

Curtin University's Embodied Conversational Agents (ECA) combine an MPEG-4 compliant Facial Animation Engine (FAE), a Text To Emotional Speech Synthesiser (TTES), and a multi-modal Dialogue Manager (DM), that accesses a Knowledge Base (KB) and outputs Virtual Human Markup Language (VHML) text which drives the TTES and FAE. A user enters a question and an animated ECA responds with a believable and affective voice and actions. However, this response to the user is normally marked up in VHML by the KB developer to produce the required facial gestures and emotional display.

A real person does not react by fixed rules but on personality, beliefs, previous experiences, and training. This thesis details the design, implementation and pilot study evaluation of an Affective Personality Model for an ECA. The thesis discusses the Email Agent system that informs a user when they have email. The system, built in Curtin's ECA environment, has personality traits of Friendliness, Extraversion and Neuroticism. A small group of participants evaluated the Email Agent system to determine the effectiveness of the implemented personality system. An analysis of the qualitative and quantitative results from questionnaires is presented.

Acknowledgement

I would like to thank my supervisors Andrew and Donald for their support and encouragement during this research. At many stages during this research I benefited from their advice. Their careful editing contributed enormously to the production of this thesis. Without them, the research would have been a lot harder and my thesis would have been provided with many more grammatical mistakes.

Thanks go to Dr Lyndall Steed, a senior lecture in the Department of Psychology, for her direction of the psychological personality theories.

Thanks also go to the ten participants of the evaluation, for their patience using the experiment application for two weeks and for providing qualitative feedback.

And finally, I would like to thank my parents, Yunbai Xiao and Xiuyun He, my brother Tianshan Xiao and my beloved Zihui Yang, for their support of me throughout this research, and for many years preceding it.

He Xiao
Perth, Western Australia
2007

Table of Contents

Chapter 1
Introduction

1.1 Research Background

Embodied conversational agents (ECAs) are synthetic lifelike characters which can interact with humans using some natural modalities of human-human communication, such as speech, facial-, hand- and body gestures (Cassell et al. 2001; Ruttkay, Dormann & Noot 2002). The Curtin Affective User Interfaces (AUI) group have extensively researched Facial Animation (Beard 2004; Marriott 2002; Marriott et al. 2001b; Marriott, Pockaj & Parker 2001a), Emotional Text-to-Speech (Stallo 2000) and Dialogue Management (Marriott 2001c; Marriott & Shortland-Jones 2003), and developed a high level markup language - VHML (VHML 2001) which can control the overall action of a virtual character. Several virtual characters have been created based on the Curtin's ECA system, such as FAQBot (Beard et al. 1999), MetaFace (Beard & Reid 2002), Virtual Sales Person (Marriott, Pockaj & Parker 2001b), Virtual Lecturer and Virtual Weather Woman (Dam & Souza 2002).

It has been proposed that personality and emotion are necessary for building believable ECAs (André et al. 1999; Barker 2003; Breese & Ball 1998). Initially, most computer science researchers studied emotion instead of personality when trying to enhance the believability of ECAs. Several emotional agents have been developed based upon the OCC model (Ortony, Clore & Collins 1988) as well as Ekman's emotion model (Ekman, Friesen & Ellsworth 1972). In contrast to emotional agents, there are few implemented personality models, although psychological personality theories do exist (Eysenck & Eysenck 1963; Leary 1957; McCrae & John 1992).

This research will design, implement and evaluate an affective personality model for an ECA (Xiao et al. 2005).

1.2 Research Objectives

The purpose of this research is to:

- Design a specification for an affective personality model based on existing research in the computer science and psychology literature.
- Convert the specification into a suitable computational model.
- Implement the model on Curtin's ECA system.
- Evaluate the model in terms of user identification of the modelled personality, the user satisfaction of the human computer interaction, and the model's ability to display emotion according to its personality.

1.3 Outline of the Thesis

The content of this thesis is structured as follows:

- Chapter 2 presents a literature review of the theories upon which this research is based.
- Chapter 3 describes issues that relate to the Research Methodologies of this research.
- Chapter 4 discusses the issues of personality modelling that involve psychological personality modelling and computational personality model.
- Chapter 5 introduces the issues that arose from implementing the computational personality model into Curtin's ECA system.
- Chapter 6 describes the data collection of this research. The three steps of the experiment are introduced in this chapter.
- Chapter 7 details the data analysis and results of the evaluation.
- Chapter 8 presents the conclusion for each hypothesis as well as indicating future directions.

Chapter 2

Research Issues

2.1 Introduction

This chapter presents a literature review of the theories upon which this research is based. The literature review is not an end in itself, but is a means to the end of "identifying the worthy research issues" (Perry 1998). It is this point that prompts the title of this chapter being "Research issues" rather than "Literature review". The theories to be discussed are, in the terms of Perry's structure of theses, *parent theories* and *research problem theories* (Perry 1998).

- The *Parent theories* come from the fields of Embodied Conversational Agents (ECAs) and personality theory. An overview of the psychological literature in relation to computer science research on emotion, mood and personality is included in this section.
- The *research problem theories* section investigates the application of personality to ECAs. The topics that this section will address are: previous ECA personality research, and the requirements for the evaluation of ECA system. This will include description of: MPEG-4 Facial Animation, Text to Emotional Speech (TTES), Virtual Human Markup Language (VHML) and Dialogue Management (DM). This section will also discuss the research at Curtin over the last five years in implementing and evaluating personality models.

Based on the discussions of these Research Issues, the relationship between the above *parent theories* and *research problem theories* that were used in this thesis will finally be presented in detail in Figure 2.12.

2.2 Parent Theories

2.2.1 Embodied Conversational Agents

Embodied Conversational Agents (ECAs) or Talking Heads (THs) are synthetic lifelike characters which can interact with humans using some natural modalities of human-human communication, such as speech, facial-, hand- and body gestures (Cassell et al. 2001). This section will introduce research aspects of ECAs that are related to personality: physical appearance, conversational and non-verbal behaviours, consistency and believability. It will be shown later in the thesis that the *physical appearance* and *conversational and non-verbal behaviours* are two aspects that are controlled by the personality model. *Consistency* and *believability* are two targets that the personality model also wants to achieve. A suitable *evaluation methodology* will also be discussed.

2.2.1.1 Physical appearance

Table 2.1 shows design aspects of the physical appearance of ECAs that has been proposed by Ruttkay et al (2002).

Table 2.1: Physical Appearance of ECAs (Ruttkay, Dormann & Noot 2002)

Physical Aspects	**Categories**
Personification	Is the ECA: • Human-like (Profession, age and gender etc) • Living creature (Dog, cow, ant etc) • Non-living object (Microsoft's paper clip)
Realism	Is the model: • Realistic • Artistic • Exaggerated cartoon-like
Dimensions	Is the model: • 2D model • 2D 'cut-out' model, which can change orientation • 3D model
Physical details	What parts of the body are covered by the model: • Head • Head and neck • Torso • Full body What details are given: • Mouth • Hands

Physical appearance can influence the perception of an agents' personality by the user. Dryer (1999) found that rounder shapes and bigger faces are perceived by humans as extraverted and agreeable, whilst bold colors and big bodies are perceived as extraverted and disagreeable.

Different categories have different application domains. For example, with regards to gender, Melson (2000) indicated that female agents tend to perform help-related tasks such as guiding the users through a new program, whilst male agents tend to perform tasks like solving homework problems. The work of Wonisch and Cooper (2002) focused on the user preference of various appearance characteristics of agents. They found that the dominant factor determining preference is the contextual appropriateness of the agent with the content area. For example Microsoft's paper clip agent is easily recognizable with a relationship to the tasks of Microsoft Office, and a doctor agent is preferred to assists the user in medical issues.

Thus, physical appearance should be carefully chosen since it may influence a user's attitude to ECAs (Isbister 1995). However, the developer should not always choose the ECA's appearance. It should be able to be changed by the ECA itself during the interaction. Design appearance is a behaviour that can reflect some personality traits. Therefore any developed ECA should be able to change its hairstyle, lip colour, beard style or even wear glasses and jewellery so as to reflect its personality. However, some fixed appearance characteristics of an ECA, such as gender and head shape, cannot be changed otherwise a user will perceive it as a new ECA.

2.2.1.2 Conversational and Non-verbal Behaviours

Conversation is the major behavioural aspect for the communication between ECAs and humans. *Conversational behaviour* allows three steps:

- A conversational agent should be able to know what the user said either by allowing the user to input text directly or by speech recognition.
- A conversational agent should know what it wants to say and in what style.

A conversational agent can output what it wants to say either by text or by speech. *Non-verbal behaviour*, such as facial display, or hand and body gestures, is necessary to reinforce the verbal conversation of ECAs (Melson 2000). For example, the user may prefer ECAs to nod their head when agreeing or to cross their arms when upset. In addition, ECAs will be more believable when emotion is added to conversational and non-verbal behaviours(André et al. 1999; Barker 2003; Bartneck 2002;

Kshirsagar 2002). An ECA with emotional speech is more believable than an ECA with normal speech (Stallo 2000).
Conversational and non-verbal behaviours will be discussed further in the personality-modelling chapter since the computational model needs to control these behaviours to achieve consistency. These behaviours involve some techniques that are currently under research, such as the Facial Animation Engine (FAE), Text to Emotional Speech (TTES), Dialogue Management (DM) and Speech Recognition (SR). Those techniques that are related to the implementation of this research will be described in the section on research problem theories. Therefore, any developed ECA system should be able to control its Conversational and non-verbal behaviour.

2.2.1.3 Consistency

Users like the ECA more when it is consistent in its attitude and responses (Cassell et al. 2001). This means that a user prefers an ECA that always has specific feelings about specific events or objects, always has the same conversation style, and does not change its beliefs or goals for a long time (Cassell et al. 2001). A stable pattern of feelings, thoughts and behaviour is needed to make ECAs consistent and this pattern can be described as the personality of the ECA. Feelings, thoughts and behaviour are the three main aspects of the personality model and these aspects will be discussed throughout this thesis. In addition, personality is also an important factor for the believability of an ECA. Therefore, any developed ECA system should exhibit consistency.

2.2.1.4 Believability

Believability is the ability for an ECA to convey the “illusion of life” to the user (Isbister & Doyle 2002). So an ECA is believable if it acts according to the expectations of the user (Ruttkay, Dormann & Noot 2002). Emotion and personality are necessary to make an ECA believable since these two factors are expected by a user when the user interacts with an ECA (André et al. 1999; Bartneck 2002; Kshirsagar 2002). Appropriate evaluation must be done in order to know whether an ECA is believable or whether an emotion or personality model can make an ECA more believable. Believability should not be equated to realistic, and believable agents do not always have a 3D realistic head. Cartoon like agents, such as Bates’s

Woggles (Bates 1994) and Beard's Spiky Boy (Beard 2004) have been proven to be believable. Therefore, any developed ECA system should be believable.

2.2.1.5 Affective ECAs

For an ECA to be affective it will have and express emotions (Picard 1997). A user of an affective ECA system should find the ECA believable, find the interaction satisfying, and engage with the system. An Affective Computing Portal can be found at http://www.bartneck.de/link/affective_portal.html.
Therefore, any developed ECA system should be affective.

2.2.1.6 Evaluating ECAs

Table 2.2: Evaluating ECAs (Ruttkay, Dormann & Noot 2002)

Evaluation	Categories
Evaluation Methodology	What to compare an ECA to? • Traditional systems with similar task • Other ECAs • Human
	Testing by what users? • Age, Gender, Ethnicity and Personality traits etc.
	Data collection: • Observation, Experiment, Survey and online survey, Questionnaire, Interview, Usage data etc.
Evaluation of usability of ECAs	• Learn ability, Efficiency, Errors etc.
Evaluation of user perception of ECAs	• Helpfulness, User's satisfaction, Believability, Trust and Engagement etc.

Ruttkay et al (2002) categorised various evaluation methodologies, evaluation aspects of usability, and user perception of ECAs which are normally used for evaluating ECAs (see Table 2.2). Rousseau and Hayes-Roth (1997a) designed experiments to evaluate their social psychological model in terms of identification of a personality and believability of the agents. Their results showed that users are usually able to correctly recognize personalities, and believe in agents who have a consistent behaviour (Rousseau & Hayes-Roth 1997a). However, Rousseau's evaluation only had eight participants and the scenario was not long enough to enable users to observe changes of moods. Nass, Isbister and Lee (2000) designed a more complex experiment to evaluate personality. A few weeks before the experiment ran, every participant was asked to fill in some psychology forms and participants were

classified as either introverted or extroverted (Nass, Isbister & Lee 2000). Data was collected from twenty participants from the introverted group and twenty from the extroverted group. These standardised classification techniques were seen as a sound methodology for conducting personality research, and it was seen that the user in identifying the personality of the ECA could also use these forms.
Nass's results proved that participants were able to identify both verbal and non-verbal personality cues. Some of the methodologies and aspects in Table 2.2 such as online survey, questionnaire, believability and engagement have been adopted by this research. Rousseau and Nass's evaluation methodologies were also studied for the evaluation of this research. The evaluation of the personality model will be described in detail in the research methodology chapter, and in the evaluation and data analysis chapter.

In summary therefore, any developed ECA should have an appropriate alterable physical appearance, and be equipped with various conversational and non-verbal behaviours. It should be believable, affective and consistent in its attitude and responses. It also needs to be evaluated in one or more human computer interaction scenarios using effective evaluation methodologies.

2.2.2 Personality Theories

2.2.2.1 Definition

Personality is explained by people every time they answer the question: "What is she or he like?" (Hewstone, Fincham & Foster 2005). Personality represents those characteristics of the individual that account for consistent patterns of thinking, feeling and acting (Pervin & John 2001). This broad definition of personality also indicated that personality is what makes people different from one another (Hewstone, Fincham & Foster 2005). The definition can be adopted to describe the personality of an ECA as a relatively stable pattern that affects feelings, thoughts and behaviour, and differentiates it from another ECA. Since psychologists have been interested in personality since the beginning of the last century, many computer science researchers (Kshirsagar 2002; Nass et al. 1995; Rousseau 1996) have used the psychology literature for modelling an ECA's personality.

2.2.2.2 Psychological Personality Theories

Personality theory is a major branch of psychology and a number of approaches have been developed during the last century:

- The first comprehensive theory of personality was Freud's *psychoanalytical theory* (Freud 1933) that attributed an individual's thoughts and actions to the unconscious in regulating behaviour.
- Allport (1937), McCrae and John (1992) and Eysenck and Eysenck (1963) proposed *trait theories* that used traits as descriptors to describe personality.
- *Behavioural theories* (Skinner 1953) were derived from laboratory experiments and they described human action and choices as behaviour learned and reinforced from the environment (Pervin & John 2001).
- *Biological theories* (Thompson 1989) attempted to explain differences in behaviour in terms of differences in physiology, and in particular, brain function (Hewstone, Fincham & Foster 2005).
- *Social learning theories* explained differences in behaviour in terms of a continuous reciprocal interaction between cognitive, behavioural, and environmental determinants (Bandura 1977).
- *Social cognitive theories* stemmed from social learning theories and are similar to behavioural theories. Social cognitive theorists focus on how a human and the environment interact (Myers 2004).

Trait theories and social cognitive theories may provide a useful basis for the computational personality model:

- Trait theories are widely used by computer science researchers to measure the personality characteristics that they have modelled, and the perceived personality of the agent (see Table 2.3). Many of the computational models of personality are based on trait theories because the traits that originate in everyday language are easy to understand, and the conversion from trait dimensions to a computational model is simple (Gulland 2001).
- The psychology theories of social learning were studied by computer science researchers so as to model personality (e.g. the personality model of Rousseau and Hayes-Roth (1997b)). The theories of social cognition stemmed from social learning theories and also adopted some behavioural theories. Therefore, social cognitive theories are new theories and will be useful for personality modelling.

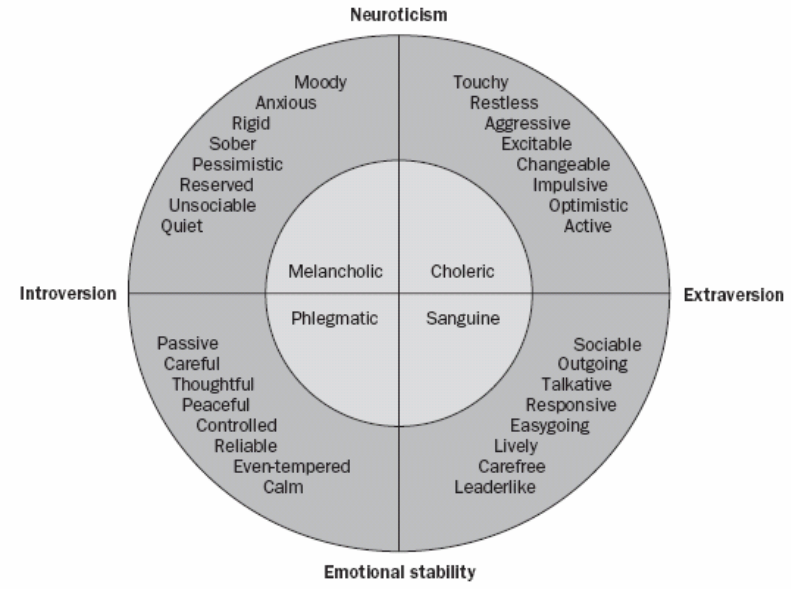
2.2.2.3 Trait Theories

Traits are labels given to consistent aspects of personality, and are viewed as continuous dimensions (Hewstone, Fincham & Foster 2005). Whether all people possess the same dimensions of personality and differ only in the value is still debated by psychologists. However, most trait theorists assume that all people have a fixed number of basic dimensions of personality. For example, Eysenck and Eysenck (1963) used two basic traits of personality, Cattell used 16 dimensions and McCrae and John's (1992) developed the Five Factor Model.

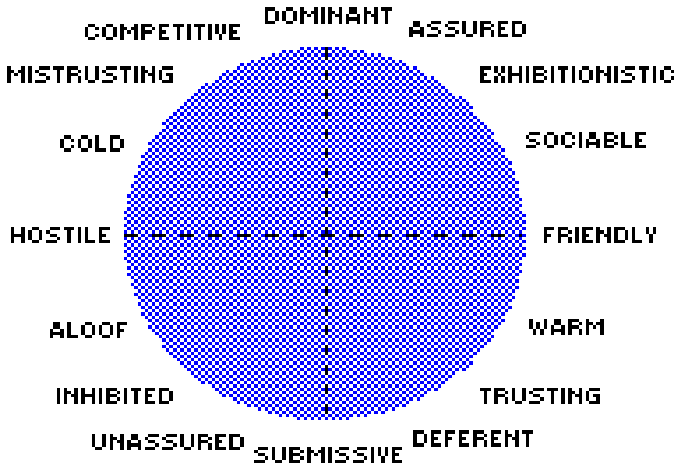
- Eysenck and Eysenck (1963) viewed the traits of *Extraversion* and *Neuroticism* as independent and believed that different personalities arose from differing combinations of these supertraits (see Figure 2.1 (a)).
- The Five Factor Model (McCrae & John 1992) determined 5 key traits in human personality: *Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism*. *Extraversion* and *Neuroticism* are defined in the same way as Eysenck's two basic traits. Among these 5 traits, *Extraversion*, *Neuroticism* and A*greeable* are more popular for modelling personality for agents since *Openness* and *Conscientiousness* are too abstract to modelling. This is confirmed by the three projects of André (2000): *Puppet* and *Inhabited Market Place* reduced the five traits to traits of *Extraversion* and *Agreeableness*, and *Presence Prototype* included traits of *Extraversion* and *Agreeableness* and *Neuroticism* (see Section 2.3.1.2 for detail). The Five Factor Model is the most popular trait theory to be adopted by computer science researchers (see Table 2.3).
- Interpersonal theory (Leary 1957) deals with a human's interaction patterns, which vary along the dimensions of *Dominance* and *Friendliness* (see Figure 2.1 (b)). Researchers (see Table 2.3) tend to use interpersonal traits for modelling an agent's personality since it can be used to describe the characteristics of the interaction between computer agents and users.

Table 2.3: Trait Theories

Trait theories	Dimensions	Computer Science Cases
Five Factor Model	Openness Conscientiousness Extraversion Agreeableness Neuroticism	· Multilayer Personality Model (Kshirsagar 2002) · Fuzzy Agents with Personality (Ghasem-Aghaee & Ören 2003) · André(1999)'s three lifelike characters project
Interpersonal theory	Dominance Friendliness	· Breese and Ball(1998)'s emotion and personality model · Nass et al.(1995) computer-based personality



(a)



(b)

Figure 2.1: (a) Eysenck's Two Traits (Eysenck & Eysenck 1963) and (b) Interpersonal Traits Theory (Leary 1957)

Not all the personality traits of an ECA come from the psychology literature. Rousseau (1996) proposed 16 dimensions of personality that are based on the processes that intelligent agents usually perform. These 16 dimensions have been successfully applied in Cybercafé and Bui's ParleE (Bui et al. 2002). The modelling of an ECA's personality is also not limited to trait theories. For example, the personality model of Rousseau and Hayes-Roth (1997b) adopted the psychology theories of social learning. The behaviour of their agents varied depending on the environment, and their past experiences with similar situations.

2.2.2.4 Social Cognitive Theory

Social cognitive theories examine consistent differences in the way people process social information (Hewstone, Fincham & Foster 2005). They focus on ways in which people interact with the environment, such as how do people interpret and respond to external events, and how do people's goals, memories and expectations influence their behaviour (Myers 2004). *Reciprocal Determinism* is an essential part of social cognitive theories. Bandura (1986) explained *Reciprocal Determinism* as the process of interaction with the environment that influences internal personal factors, and behaviour. As Figure 2.2 shows, people's thoughts and feelings about risk activities influence how their bungee-jumping friends affect their current behaviour - whether or not to learn to bungee jump (Myers 2004). The interaction between these three factors is extremely complex, and will differ based on the individual, the particular behaviour being examined, and the specific situation in which the behaviour occurs (Bandura 1989).

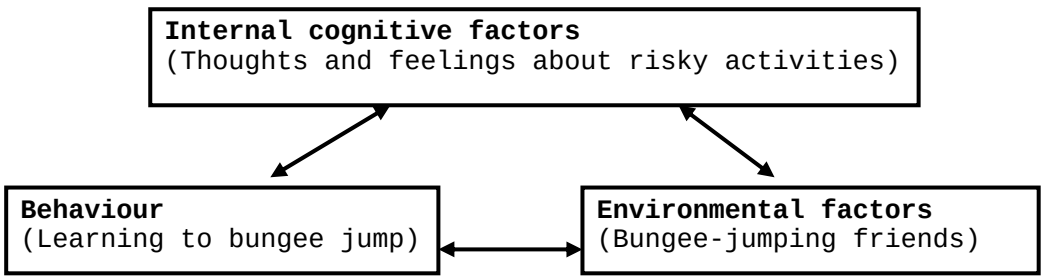


Figure 2.2: Reciprocal Determinism of Social-Cognitive Theories (Myers 2004)

One of the applications of *Reciprocal Determinism* to social cognitive theory is to understand and predict an individual's behaviour. This theory can be used to predict a

computer agent's behaviour based on the internal and environmental factors of the agent. This will be discussed in detail in the personality-modelling chapter.

2.2.3 Relations of Personality, Mood and Emotion

It has been proposed that personality, mood and emotion are necessary for building believable ECAs (André et al. 1999; Barker 2003; Bartneck 2002; Breese & Ball 1998; Egges & Kshirsagar 2004; Kshirsagar 2002). Time duration is the main difference between these three components (Kshirsagar 2002; Moffat 1997; Wilson 1999). Figure 2.3 (from Kshirsagar (2002)) shows that emotion is a momentary state and mood is more static than emotion. Personality in contrast, does not change over a long time. The vertical axis represents the three layers in Kshirsagar's Multilayer Personality Model (see Section 2.3.1.1). The high layer represents the more abstract aspects and the low layer represents the more concrete aspects of personality. Wilson (1999) suggested that these three layers have different priorities for controlling behaviour: emotions have the highest priority whilst personality has the lowest. Moffat (1997) also indicated a focus difference between emotion and personality. Emotions are focused on specific events, actions and objects whilst personality is more general. An investigation of relation and interaction among emotion, mood and personality is necessary to build a powerful personality model.

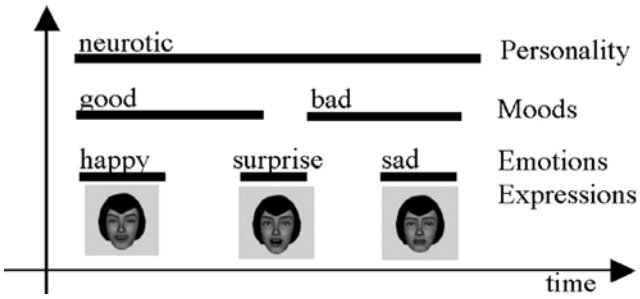


Figure 2.3: Personality, Mood and Emotion (Kshirsagar 2002)

2.2.4 Emotion Theories

ECA researchers (Bartneck 2002; Bates 1994; Vélasquez 1997) considered emotion as an essential part of a believable ECA and their research focussed on two aspects: emotional facial, body and vocal expressions, and the appraisal of an agent's emotions based on its reactions to objects, events and actions. They developed a number of emotional applications: for example, Reilly and Bates' Em (Reilly & Bates 1992), Gratch's Émile (Gratch 2000), Vélasquez's Cathexis (Vélasquez 1997) and

Prendinger's SCREAM (Prendinger, Descamps & Ishizuka 2002). Many of these applications employed the cognitive appraisal model of OCC (Ortony, Clore & Collins 1988) (see Figure 2.4).

Emotions in the OCC model are the results of three kinds of appraisals (Reilly & Bates 1992):

- The appraisal of the pleasantness of events with respect to the agent's goals.
- The appraisal of the approval of the actions of the agent or another agent with respect to a set of standards for behaviour.
- The appraisal of the liking of objects with respect to the attitudes of the agent.

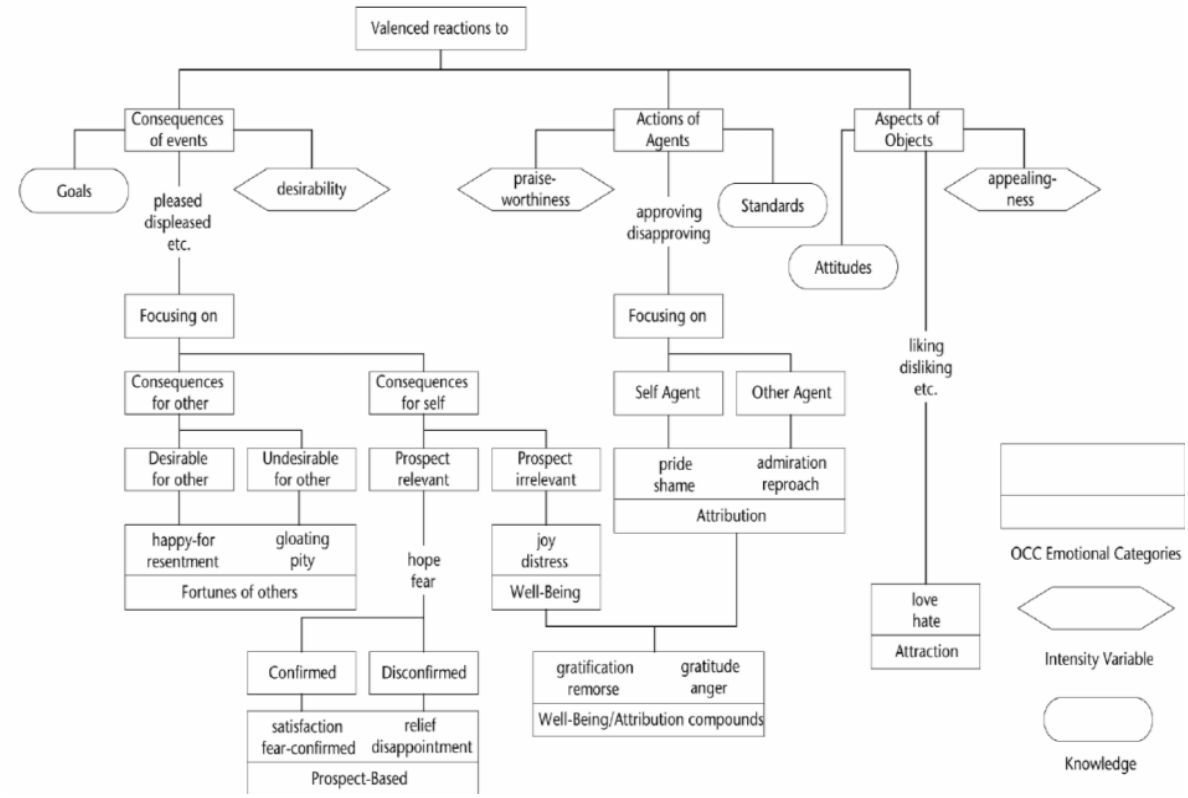


Figure 2.4: The OCC Model (Bartneck 2002; Ortony, Clore & Collins 1988)

Table 2.4: Basic Emotions of the OCC and Ekman Models

Model	Emotion types
OCC	Happy-for Resentment Gloating Pity Joy Distress Pride Shame Admiration Reproach Love Hate Satisfaction Fears-confirmed Relief Disappointment Gratification Remorse Gratitude Anger Hope and Fear
Ekman	Happiness Anger Disgust Sadness Surprise and Fear

Table 2.4 shows the emotion types of the OCC model and of Ekman's model (Ekman, Friesen & Ellsworth 1972). The 22 emotion types modelled in OCC are used for appraisal of the agent's emotions while the 6 basic emotions of Ekman's model are widely used for emotional facial and vocal expressions of an agent.

Bartneck (2002), using OCC, split the emotion process into the following phases:

- *Classification* - Find what emotional types are affected using the evaluation of an event, action or object based on the ECA's knowledge.
- *Quantification* - Calculate the intensities of the affected emotional types.
- *Interaction* – Use the intensity to update the ECA's current emotional state.
- *Mapping and Expression* - Map the 22 OCC emotion types to a possible lower number of expression emotions available to the ECA, such as Ekman's 6 emotions. The ECA then renders its emotion through facial, body and vocal expressions.

Additionally, a phase is needed to blend the emotion types for an expression. This phase solves the problem of more than one emotion type being generated at each moment. For example, the ECA may be both happy because of an interesting topic, and disgusted with the user's silly questions. The emotion-blending problem is a considerable problem for the expressing of emotions, especially for facial expressions. Ekman and Friesen (1975) proposed a model of blending facial expressions by combining the upper part of one expression with the lower part of the other one. Following Ekman's findings, Ochs et al. (2005) proposed an emotion blending model and successfully implemented it in the Greta system (Pelachaud & Bilvi 2003) for each blend type, superposition and masking (see Figure 2.5).

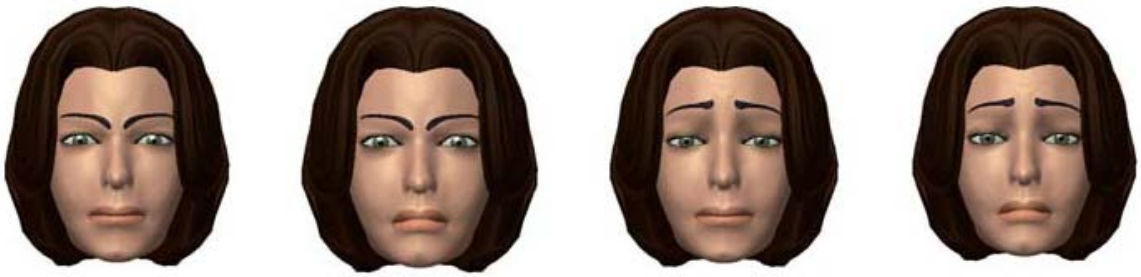


Figure 2.5: Facial Expressions of the ECA Greta (Ochs et al. 2005).

The OCC model is designed to model humans in general, and researchers intend to use the OCC model to create a “neutral” agent that behaves like a “normal” human (Bartneck 2002). However, by using a neutral agent it is difficult to satisfy different users, since users will have a different opinion of this neutral agent. The neutral emotional agent needs to be equipped with something that make its behaviour different from other agents, and caters for the preferences of different users. That is, a personality.

2.2.5 Mood Theories

Adopting a psychobiological perspective, mood is explained by Vélasquez (1997) as a low level of arousal within an emotional system, whilst emotion is explained as a high level of arousal. Mood is a necessary bridge which links personality to emotion expression since a personality models’ high level descriptions make it difficult to directly control the emotions that are visible on the virtual face (Kshirsagar 2002).

Kshirsagar (2002) simply defined mood as one dimensional, going from bad mood to good mood, while Rousseau’s social-psychological model (Rousseau & Hayes-Roth 1997b) divided mood into two categories: agent-oriented moods, and self-oriented moods. The agent-oriented moods are directed toward other individuals whereas the self-oriented moods are not (Rousseau & Hayes-Roth 1997b). The distinction between self-oriented and agent-oriented moods can make a character rather happy in general, yet remain angry at a particular individual because of what he/she did to the character (Rousseau & Hayes-Roth 1997b).

This section discussed Parent theories that come from the fields of Embodied Conversational Agents (ECAs) and personality theory. A review of the psychological literature in relation to computer science research on emotion, mood and personality is also involved. The next section will present research problem theories – issues that relate directly to this thesis.

2.3 Research Problem Theories

2.3.1 Previous Personality Researches

2.3.1.1 Kshirsagar's Multilayer Personality Model

Kshirsagar (2002) proposed a layered approach to personality modelling (see Figure 2.3). That personality model had the following three layers (Kshirsagar 2002):

- The personality layer does not change over time and it caused deliberative reaction and affected how moods changed. The Five Factor Model was adopted to describe the personality.
- Moods were affected by personality and also affected by the emotions.
- The emotions in the lowest layer link directly to expressions, and are influenced by moods and previous emotion types.

Bayesian Belief Networks were adopted in the personality model to handle the uncertainty, and gave a structured probabilistic framework to represent and calculate the complex and abstract concepts related to the three layers (Kshirsagar 2002). The model handled the uncertainty of the expression of emotions effectively by taking into account personality traits, mood and previous moods, previous emotions and the current emotion generated from the OCC appraisal (Egges & Kshirsagar 2004). The multilayer personality model was implemented in a virtual human chat system successfully. However, the limitation of the multilayer personality model is that the personality characteristics can only affect the emotional expressions of the agent. An ideal personality of a believable agent should be able to guide the behaviour, thoughts as well as emotions.

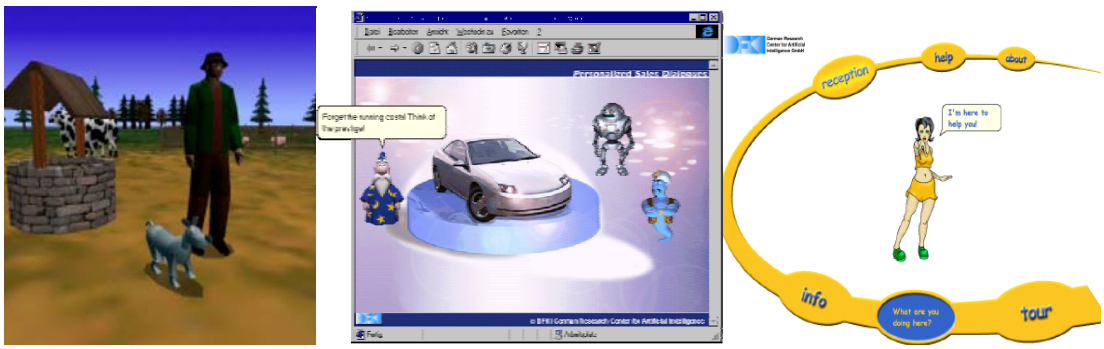
2.3.1.2 André's Three Lifelike Agents Projects

The three projects of André et al (1999) described below, adopted models of emotions and personality to control the social behaviour of interactive animated agents. These three projects used a similar approach towards modelling emotions and personality traits. The Five Factor Model of personality was used as a description of an agents' personality and the OCC model of emotions was used in the *Puppet* and *Presence* projects to determine the agents' emotional type in response to events.

- The *Puppet* project, based on Reilly's (1996) theoretical framework of learning through externalisation, developed a virtual reality environment, the

virtual puppet theatre (see Figure 2.6 (a)) as an interactive learning environment to support the development of a child's emotional intelligence skills (André et al. 2000). This project reduced the Five Factor Model to two common dimensions of extraversion and agreeableness.

- The second project used an *Inhabited Market Place* with a group of agents to communicate product information and give sales presentation (see Figure 2.6 (b)). Personality traits were used to modify the agents' virtual actor role in sales presentations. The agents not only communicated plain facts about certain subjects but also presented them from a point of view that was guided by their specific personality characteristics and emotions (André et al. 2000). Similar to the *Puppet* project, this project used the dimensions of extraversion and agreeableness in the Five Factor Model.
- The *Presence* project used agents as virtual receptionist guides for visitors. An internal model of the agent's affective state guided the conversational dialogue between agent and user (André et al. 2000). The dimensions of extraversion, agreeableness and neuroticism of the Five Factor Model were considered and the virtual receptionist was modelled as an extravert, agreeable and emotionally-balanced agent (André et al. 1999).



(a) **(b)** **(c)**

Figure 2.6: (a) Puppet 3D Environment (b) Inhabited Market Place and (c) Presence Prototype (André et al. 2000)

2.3.1.3 Nass's personality research

Psychologists Byrne et al. (1967) indicated that humans prefer to interact with others who are similar in personality to themselves. Following this finding, human computer interaction research by Nass et al. (1995) attempt to demonstrate that this similarity

attraction applied equally to the interaction between human and computer, and found that users are more satisfied with the interaction when the user and the computer have similar personalities. One of the hypotheses of Nass's research was: "There will be a cross-over interaction between the user's level of dominance and the computer's level of dominance such that users will be more favourably disposed toward the computer, and more satisfied with the interaction, when the levels of dominance / submissiveness are similar" (Nass et al. 1995). This hypothesis was proven and the results showed that submissive users preferred a submissive agent instead of a dominant one. It follows that an ECA system that allows a user to choose or develop a suitable personality for the ECA they interact with, will result in a more satisfying interaction.

2.3.1.4 August Dialogue System

The August system is a multi-modal spoken dialogue system with an animated agent called August with whom the user interacts (Gustafson, Lindberg & Lundeberg 1999). As Figure 2.7 shows, the agent August was made to look like the 19th century Swedish author August Strindberg, who was famous for making well-known statements about politics, women and reviewers (Lundeberg & Beskow 2005). This was the first system that attempted to give an agent some actual person's personality characteristics. Behaviour was taken into account as part of the personality; the agent had the behaviour of twisting and stretching his beard and moustache (Lundeberg & Beskow 2005). Ekman's 6 basic emotions were adopted to display the agent's facial expression. However, the August Dialogue System doesn't have an underlying Personality Model to control the agent's emotions and behaviour.

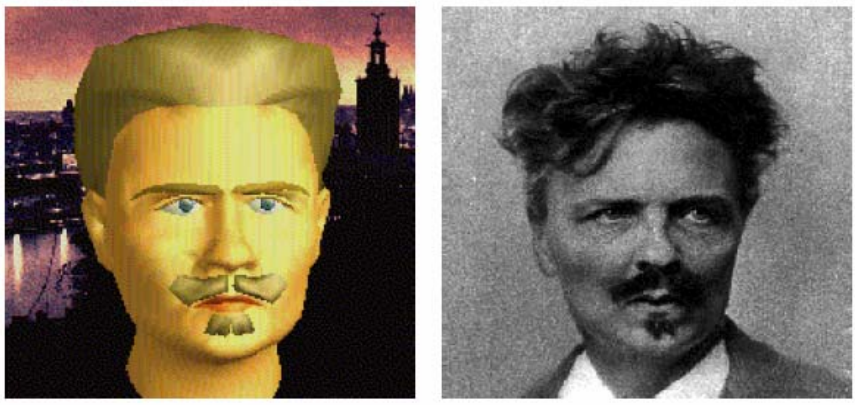


Figure 2.7: The ECA August and the 19th Century Swedish Author August Strindberg (Lundeberg & Beskow 2005).

2.3.2 An Existing Emotional ECA System

The personality model of this thesis was implemented in Curtin's ECA system (Beard 2004; Marriott 2002; Marriott et al. 2001d) to control an ECA's emotional expression and behaviour. Curtin's ECA system combines an MPEG-4 compliant Facial Animation Engine (FAE), a Text To Emotional Speech Synthesiser (TTES), a multi-modal Dialogue Manager (DM) that accesses a Knowledge Base (KB) and outputs Virtual Human Markup Language (VHML) text which drives the TTES and FAE. Several virtual characters have been created based on the system, such as FAQBot (Beard et al. 1999), MetaFace (Beard & Reid 2002), Virtual Sales Person (Marriott to be published), Virtual Lecturer and Virtual Weather Woman (Dam & Souza 2002).

The implementation of a personality model within this system was proposed so as to be able to guide facial animation, gestures and the sound of the voice of an ECA when it communicated with the user, according to its personality characteristics. The following techniques, MPEG-4 facial animation, text to emotional speech, virtual human markup language and dialogue management, will be discussed since they relate to the computational personality modelling, implementation and evaluation of this research.

2.3.2.1 MPEG-4 Facial Animation

An historical approach to facial animation was to determine a set of parameters to control the animation of a face model. This approach is often called parameterized facial animation. Parke's facial animation system, described in Parke & Waters (1996) was the earliest parameterized face animation systems. Although many facial animation systems have been developed after this initial research, most of them suffer from the same limitation: each of them uses a proprietary architecture and syntax for animating a synthetic face (Marriott et al. 2001b). A standard parameterization model is needed to make any standard compliant face model be animatable by any standard compliant player (Marriott et al. 2001b). The ISO/IEC international standard of MPEG-4 (ISO/IEC 1998), overcomes the limitation by defining a standard for the efficient representation and transportation of face media. Basing facial animation on the MPEG-4 standard also allows different frameworks to work cohesively together (Beard 2004).

Curtin's ECA system is based on an MPEG-4 facial animation system that uses the Facial Animation Engine of Lavagetto and Pockaj (1999). However, it is proposed that the personality model of this thesis does not control the low-level facial animation directly. The facial animation and text to emotional speech synthesiser is controlled by a higher level scripting language called VHML (see 2.3.2.3).

2.3.2.2 Text to Emotional Speech Synthesiser

Stallo (2000) developed a system that could add simulated emotion effects to synthetic speech, and integrated the system within the text-to-speech (TTS) module of an ECA. The system was proved to be effective by analysis of the perceived ECA emotions by listeners, and it is used by Curtin's ECA system as text to emotional speech synthesiser.

Table 2.5: Listener Responses for Utterances Containing Emotive Text with Vocal Emotion (Stallo 2000)

STIMULUS \ PERCEIVED EMOTION	Happy	Sad	Angry	Neutral	Surprised	Disgusted	Other
Happy	66.7%	4.4%	0.0%	13.3%	4.4%	2.2%	8.9%
Sad	0.0%	62.2%	4.4%	24.4%	0.0%	0.0%	8.9%
Angry	0.0%	0.0%	77.8%	1.1%	0.0%	15.6%	5.6%
Neutral	6.7%	2.2%	0.0%	71.1%	13.3%	4.4%	2.2%

Table 2.6: Percentage Of Listeners Whose Emotion Recognition Improved / Deterioration with the Addition of Vocal Emotion Effects for Emotive Text (Stallo 2000)

Emotion	Improvement	Deterioration
Happy	**57.8%**	4.4%
Sad	31.1%	17.8%
Angry	41.1%	6.7%

The summary of that research in Table 2.5 shows that users could recognise specific emotions added to a synthetic voice. The following observations can be made from the data:

- Table 2.5 shows a strong average recognition rate for all simulated emotions (seen in the high values down the diagonal line, matching row and column names).
- Confusion continued to occur for emotions, but was not as significant as other utterance types.

Both of the above observations indicate that the simulated emotions were perceived correctly without being confused with other emotions.

Table 2.6 shows the percentage improvement in recognition for untagged vs. speech markup language (SML) tagged text. All emotions received a significant increase in emotion recognition once vocal emotion was added to the utterance, with the greatest improvement occurring for happiness (57.8%) (Stallo 2000).

This text to emotional speech synthesiser is controlled by the Speech Markup Language (SML), which is a sub language of VHML (VHML 2001). The computational personality model does not control Stallo's system directly since it is low level. It is proposed that the personality model of this thesis controls the ECAs' emotional voice by generating appropriate VHML tags according to the ECAs' personality characteristics and current emotions.

2.3.2.3 Virtual Human Markup Language

A number of scripting language, such as MPML (Prendinger, Descamps & Ishizuka 2002) and GESTYLE (Ruttkay & Noot 2005), have been developed by ECA researchers to achieve a higher level of control over an ECA. The use of scripting language allows the ECA to be used in different information domains by altering scripts instead of the application code or framework (Beard 2004). The Virtual Human Markup Language (VHML) is one of the scripting languages which is based on XML/XSL and consists of the following sub languages (VHML 2001):

- EML Emotion Markup Language
- GML Gesture Markup Language
- SML Speech Markup Language

- FAML Facial Animation Markup Language
- BAML Body Animation Markup Language
- XHTML eXtensible HyperText Markup Language (only a subset is used)
- DMML Dialogue Management Markup Language

The Virtual Human Markup Language (VHML) is designed to accommodate the various aspects of Human Computer Interaction (HCI) with regards to facial animation, body animation, dialogue manager interaction, text-to-speech production, emotional representation and hyper and multimedia information (VHML 2001). VHML can be divided into three levels (see Figure 2.8) (VHML 2001). There are only five elements at the top level. At the middle level are the two sub languages, EML and GML, which control emotions and gestures. Their elements are inherited by three low level languages: SML, FAML and BAML. Apart from these three languages, there are two additional sub languages at the low level, DMML and XHTML, that can be ignored for this research. The dotted lines indicate that the languages on the lower level inherit the elements from the language on the upper level (VHML 2001).

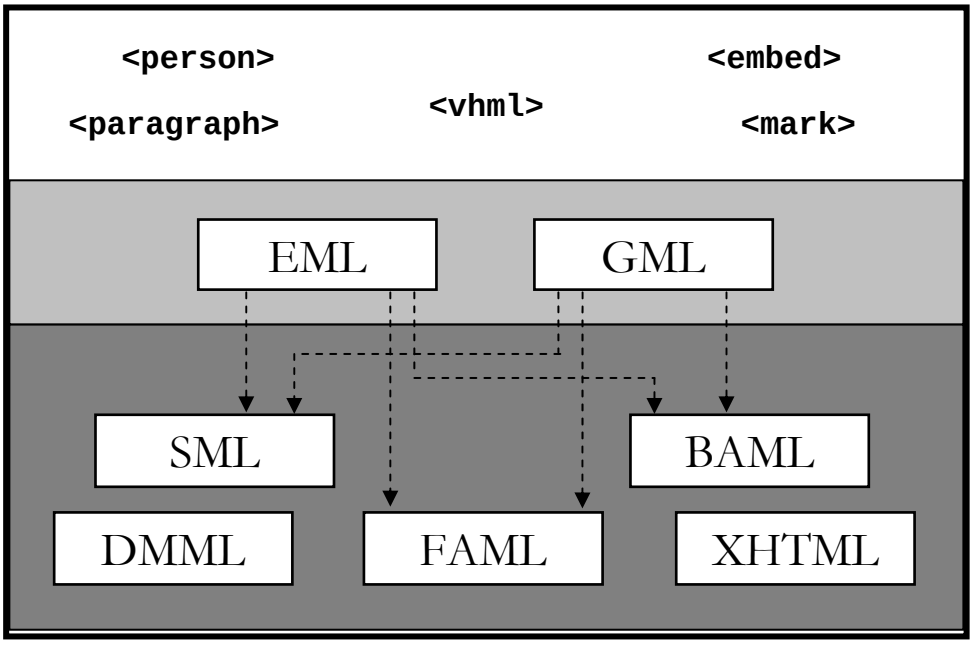


Figure 2.8: The VHML Language Structure (VHML 2001)

```
<vhml>
  <paragraph> <sad>
      <pitch range="+150%" middle="-10%">
        And here's the latest <pitch middle="-18%">news.</pitch>
        <pause length="short"/> Detectives investigating the
        <emph>brutal</emph> murder of Sarah Payne, <blink/> have
        received 200 fresh calls from the public.
      </pitch><smile/>
  </sad> </paragraph>
</vhml>
```

Figure 2.9: VHML News Item of a Newsreader ECA

The VHML example shown in Figure 2.9 controls how the news should be read by a news reporter ECA: overall a sad voice and posture, a pause after the first sentence for the listener to pay attention, emphasis on the word “brutal”, a short blink at the sentence break, and a smile at the end of the second sentence since it probably indicates good news. However, it must be noted that the marking up of this dynamic and constantly changing “news” information by the scriptwriter is tedious and time consuming. It would be more appropriate for the ECA to read the news with its own thoughts emphasis, gestures, and feelings. That is, the ECA would filter this information through its own personality by marking up the plain text with VHML (Xiao et al. 2005). The lack of personality was indicated as a considerable problem of VHML since a Virtual Human needs a personality (Marriott & Stallo 2002).

2.3.2.4 Dialogue Management

A Dialogue Management System (DMS) is a mechanism for natural language communication allowing multimodal input and is responsible for understanding what the user wants when they interact with the system (Marriott, Pockaj & Parker 2001a). The Mentor System (see Figure 2.10) is a Java-based client-server DMS designed to be a learning assistant to help students in their university studies (Marriott 2005). This system uses Perl-5 Regular Expression pattern matching for Natural Language Parsing along with a state based Dialogue Manager and a Knowledge Base marked up using VHML (Marriott, Pockaj & Parker 2001a).

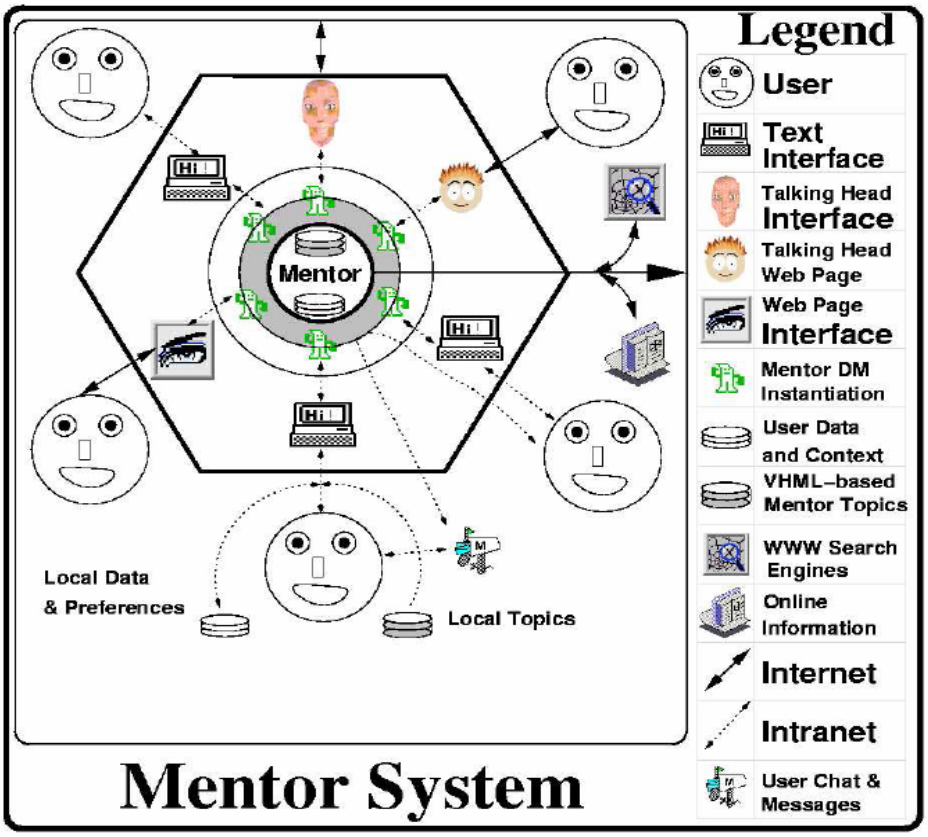


Figure 2.10: The Mentor System (Marriott 2005)

A user enters a question and an animated ECA can respond with a believable and affective voice and actions. The response to the user is generated by the Mentor System according to the topic that the user's input has matched against. For example, if a user simply types "1 + 3", one of the Mentor topics will match it with the pattern ".*\\b((pi)|(e)|([^A-Za-z]+[0-9^+\\-*/%().]+([eE][+-]?[0-9]+)?))+\\b.*" and the response is generated randomly from the possible arithmetic responses such as "I think the answer is 4" or "I have better things to do than work out maths for you." etc.

Similarly, if the user types in "what is the weather forecast", the system will give a predefined non-informative answer or will data mine a meteorological site to get the exact forecast.

Of interest, the two responses to the "1 + 3" request could represent the different personality traits or different emotions of the ECA. It would be a more believable interaction if a ECAs' personality and emotion could be considered when choosing the responses. That is, the ECA may tend to choose the second answer when he/she is hostile or he/she feels angry.

2.3.3 Curtin's Implemented Personality Applications

Shepherdson (2000) focused on specific personality traits that could be modelled in an MPEG-4 compliant Talking Head (TH): dominance, submissiveness, friendliness and unfriendliness. These traits are important in interpersonal communication. The study used facial expressions of emotions, eye behaviour (gaze behaviour, eyelid closure and openness, eyelid blinks), mouth gestures, head movements (head-turning, head-nodding) and gestures (eyebrow raising and lowering) to convey different personality traits. Each of the chosen personality traits was modelled and stored in a personality file that directed the display and controlled facial expressions.

Shepherdson (2000) had three hypotheses - the TH implementing a personality will:

- be able to communicate the spoken information to the user more effectively,
- correctly exhibit personalities as intended by the author,
- be perceived by users as a more humane interface.

Only the second hypothesis was proven conclusively although anecdotal evidence indicated some support for the other two. Also, to better differentiate the personalities, a bigger set of gestures was required and, to be an effective communicator, it was suggested that a TH required clear and audible speech from a better speech synthesiser. Further, the research concluded that since personality is not just a single value (e.g. dominant), but a continuum of values, future work should allow the blending of personality traits.

Gulland (2001) evaluated a case study built on the Five Factor Model. The Idolum framework demonstrated an idle behaviour of moods and emotions controlled by a consistent personality. In order to be more believable, Idolum took into account aspects of personality and mood, and stimuli elements from psychological models such as a time cycle (winter/summer), the weather, or a manic/depressive cycle that can affect emotional behaviour (see Figure 2.11). To avoid predictable and repetitive actions that can hinder the believability of a character (Lester & Stone 1997) , it was seen as important to incorporate a small random factor in the change of emotional levels and the relevant behaviours calculated from these levels.

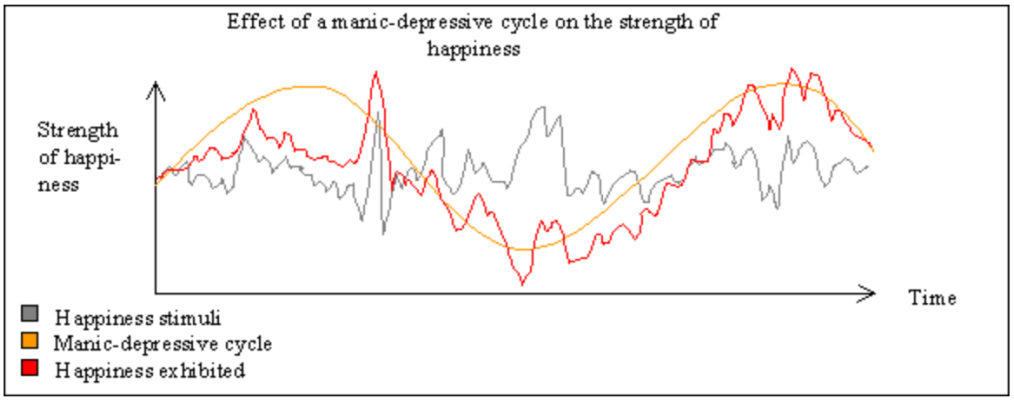


Figure 2.11: Contrived Example of the Effect of a Manic/Depressive Cycle on the Strength of "Happiness" Exhibited (Gulland 2001).

Building on this, Dam and de Souza (2002) reported on a TH system that "learnt" the user's opinions about weather conditions so as to mimic the user's likes and dislikes. Although this "mimic personality" research was done in a limited domain, the results supported the research on personality compatibility done by Nass et al (1995). Beard and Reid (2002; 2001) and Marriott (2004) continued in this research area by addressing the issues of personalities in THs, and the issues of human-TH interaction concerned with personality. For example, a user may be able to determine the personality of a TH through its voice and actions, but the opposite – the TH determining the personality of the user - is not true due to the single-modal input to the TH system from the user – plain text. To this end, the existing DM has been modified to cater for multi-modal input (such as emotion values). Although convoluted, the user can now also include emotion in the dialogue with a TH by manipulating graphical slider values for emotions. These emotions may enable a TH to determine the personality of the user over many interactions.

2.4 Conclusion

The previous sections have introduced a review of the theories upon which this research is based. As Figure 2.12 shows, the theories were divided into *parent* theories and *research problem* theories. The personality model was developed based on the theories of emotion, mood and personality and on previous successful personality research. Key aspects and theories of the ECAs in the *parent* theories section related closely with techniques of MPEG-4 FAE, DM, TTES and VHML which were then detailed in the *research problem* theories section. The above ECA

theories and techniques are investigated and used throughout the research, from personality modelling, implementation, and to evaluation. Curtin's implemented personality applications over the last five years were reviewed as well as their effect on the proposed modelling of the personality and implementation of the computational model.

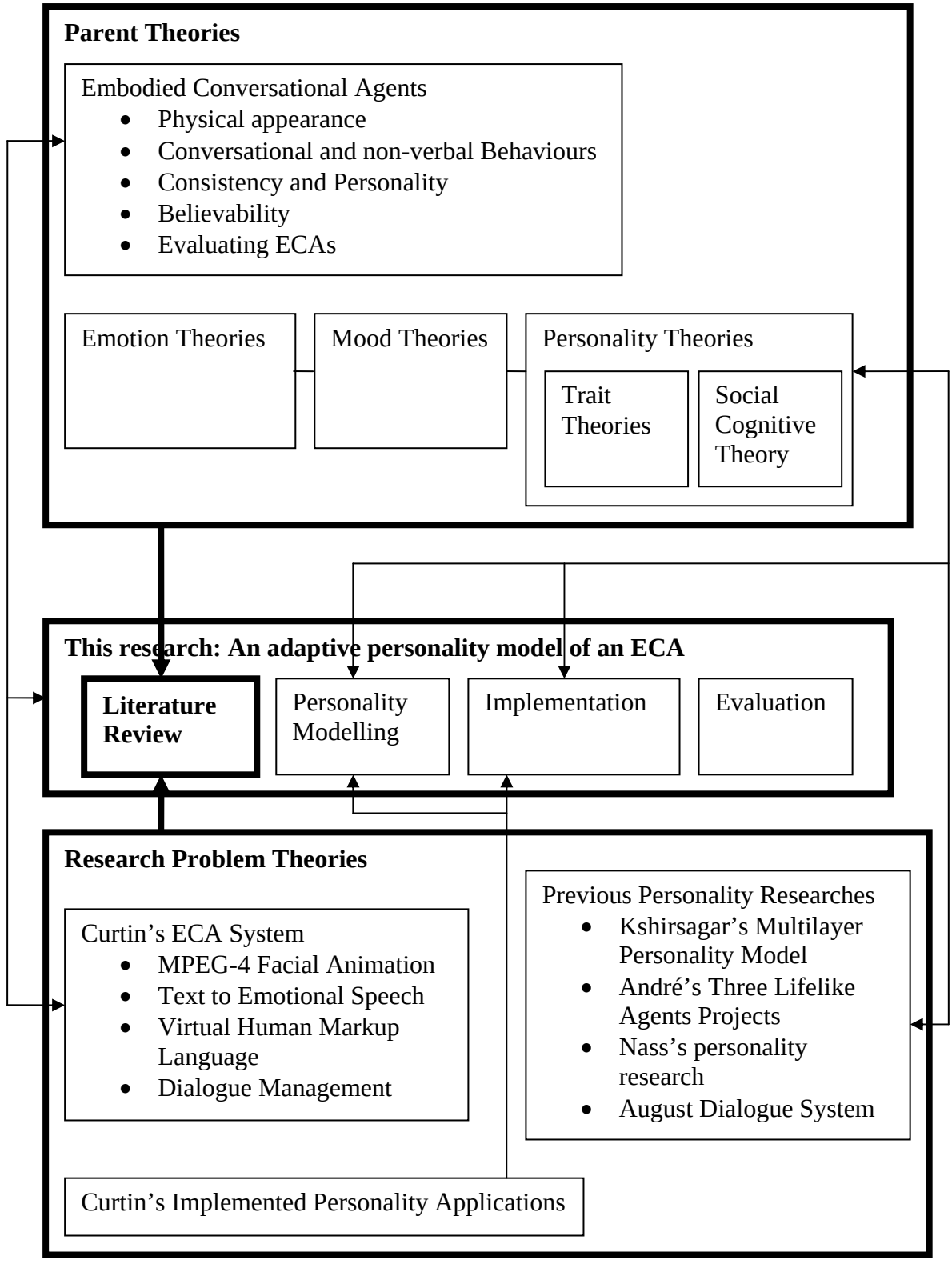


Figure 2.12: Literature Review of this Research

Chapter 3
Research Methodology

3.1 Introduction

The previous chapter presented a review of the parent theories and research problem theories of this research. In order to develop and evaluate the personality model correctly, the appropriate research methodologies (RM) had to be used. This chapter discusses issues that relate to the RM of this research. The research hypotheses that guided the development and the evaluation of the personality model are introduced. Four research objectives are also indicated and these reflect the four phases of research: design and development of the personality model, implementation and evaluation. The limitations of the research are described so that the results could be understood in their proper context. Then two types of RM are introduced: the *design and implement* methodology and the *evaluation* methodology. In addition, data collection and storage, significance and ethical considerations are also discussed.

3.2 Hypotheses

The following hypotheses are the basis for this research:

Hypothesis: A personality model will enhance the believability of an ECA and make human computer interaction more satisfying.

The research proposed to address the two sub-hypotheses:

Sub-hypothesis 1: An ECA equipped with a personality model can express emotion according to its personality characteristics.

Sub-hypothesis 2: A user can identify the personality or the personality characteristics of the ECA.

The hypothesis indicates that the overall target of this research is in enhancing the ECA believability and making the user interaction more satisfying. The first sub-

hypothesis indicates whether the personality model is affective according to its personality characteristics, and the second sub-hypothesis reflects the accuracy of the modelled personality. If the modelled personality is accurate, then it will become possible to correctly tailor the ECA personality to a particular user, and hence will benefit the interaction (Byrne, Griffitt & Stefaniak 1967) (Nass et al.).

The first sub-hypothesis is on an application level whilst the second one is on a user perception level. These two sub-hypotheses are necessary to prove or disprove the main hypothesis. All these hypotheses will be evaluated in this research using the evaluation methodologies described in Section 3.5.2. Chapter 7 will analyse the evaluation data to confirm or deny the hypotheses.

The personality model is proposed to be affective since the model will have and express emotions (Picard 1998). The emotion appraisal module (see Section 4.3.1.1) and the emotional behaviour modules (see Section 4.3.3.1) of the personality model will give the ECAs emotions and the ability to express these emotions according to its personality characteristics. The evaluation of the main **Hypothesis** and specifically **Sub-hypothesis 1** will determine whether or not the personality model is affective. Additionally, an affective ECA can interact naturally with the user and this will be tested by the evaluation of the believability.

3.3 From Action Research to Research Objectives

Action Research is a group of research methodologies, the essence of which is “learning by doing" - a group of people identifies a problem, does something to resolve it, sees how successful their efforts were, and, if not satisfied, tries again (O'Brien 1998). Kemmis and McTaggart (1988) developed a simple model of the cyclical nature of the typical action research process and each cycle has four steps: plan, act, observe and reflect (see Figure 3.1).

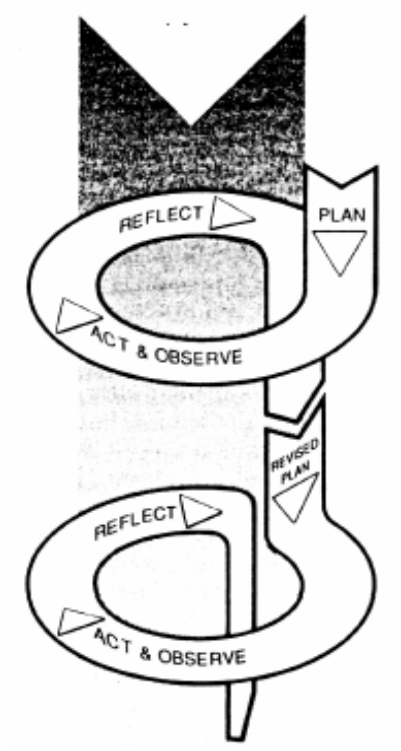


Figure 3.1: The Action Research Spiral (Kemmis & McTaggart 1988)

A limitation of the action research spiral is that it is intended for long-term research. As this research is of one-year duration, only one big cycle is possible. Figure 3.2 outlines the intended cycle. However, some micro cycles are involved in this cycle: for example, micro evaluation and subsequent reflection by the researchers was done to refine the model during the period of design and implementation of the personality model.

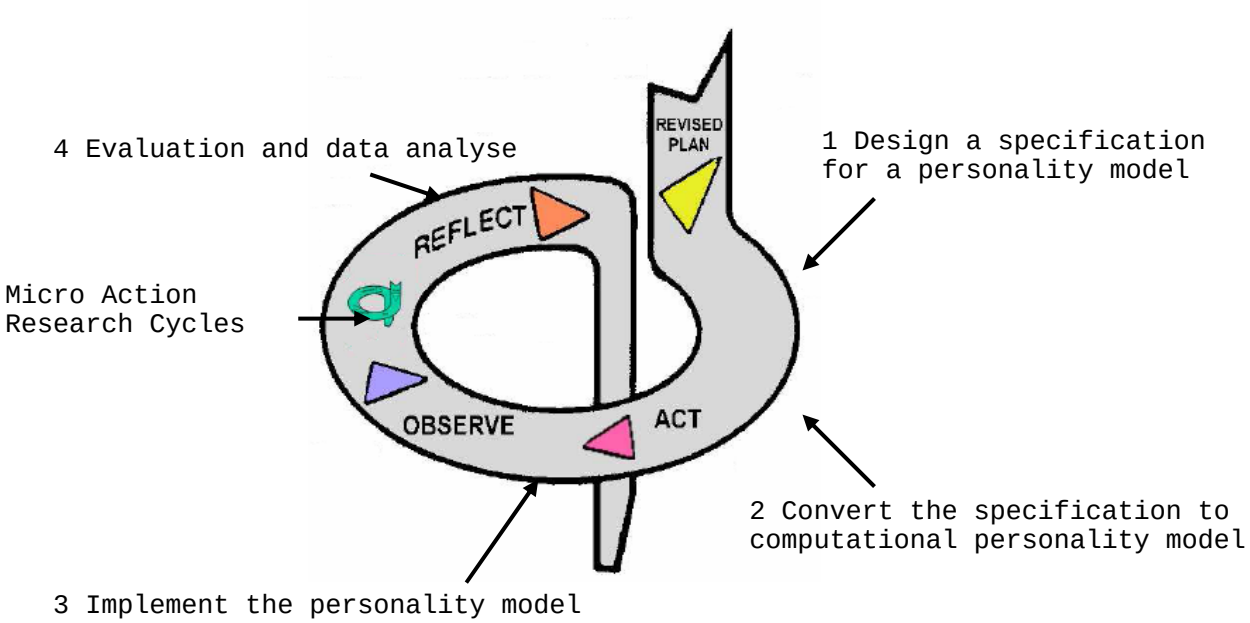


Figure 3.2: Generate Research Objectives Based on the Action Research Spiral

The literature review phase of this research identified research issues and relevant research problems concerning personality models in ECAs. Based on the existing ECA environment, the objectives of this research were:

1. Design a specification for a personality model based on existing research in the computer science and psychology literature.
2. Convert the specification into a suitable computational model.
3. Implement the model within Curtin's ECA system.
4. Evaluate the model in terms of user identification of the modelled personality, the user satisfaction of the human computer interaction and the model's ability to display emotion according to its personality.

In order to achieve these objectives, relevant research methodologies were adopted, and these methodologies can be divided into two categories: *design and implement* methodologies in the first three phases and *evaluation* methodologies in the last phase.

3.4 Limitations

The personality model was designed to provide an ECA with a personality pattern that can guide the ECA's feelings, thoughts and behaviour. However, it is unlikely that the model will be able to make this pattern as complex as real person's personality. The designed pattern was a simulation of some typical personality characteristics from the psychology literature, and the control of personality over feelings, thoughts and behaviour (Pervin & John 2001). Hence the scope of the personality for the ECA is limited. The personality model and implementation has concentrated on the traits of *Friendliness, Extraversion* and *Neuroticism* (see Section 4.2.1). These have been used in previous successful ECA personality research (André et al. 1999; Barker 2003; Bartneck 2002; Breese & Ball 1998; Egges & Kshirsagar 2004; Kshirsagar 2002).

The personality model was implemented within Curtin's ECA system and hence is subject to the limitations of that environment. Firstly, the ECA's facial animation, speech and gestures were controlled by VHML scripts, and the behaviour of the ECA was limited by the scope of VHML. Secondly, the ECA used only the face instead of the whole body. See Chapter 4 and 5 for details of the implemented personality model.

These two limitations made some important characteristics of personality difficult to implement. For example, psychologists have found that a person with a typical extravert personality will “keep moving” (Eysenck & Eysenck 1963). However, this behaviour can only be implemented as “keep moving it’s head” since the agent only is a talking head. Moving an ungrounded head may have a different effect compared to moving a head with neck and shoulder (see Figure 3.3). The Sarah model of Figure 3.3 can move its head whilst the shoulders remain stationary. The Metaface model of Figure 3.3 would simply have has its head bobbing around.

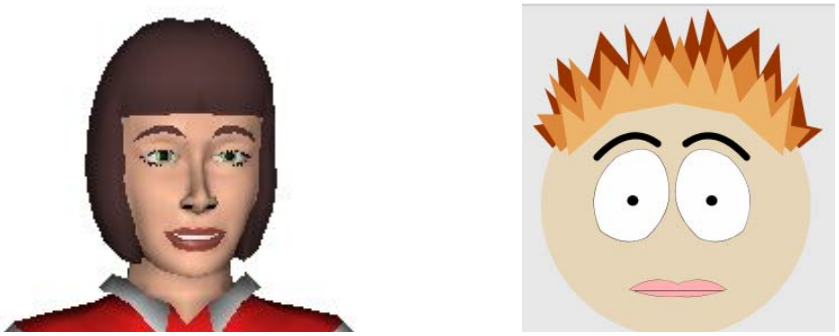


Figure 3.3 Sarah Model (Eptamedia 2005) versus existing Metaface Model

Thirdly, the personality model was in a restricted domain of interaction. The interaction of Curtin’s ECA Dialogue Management system is that a user enters a question and an animated TH responds with an affective voice and actions. This is significantly different from real human-human interaction. However, the DM is pro-active in that it can initiate dialogue. For example, it can comment on some event that occurs (such as email arriving), or simply make a random observation.

Finally, the evaluation of personality with subjects is difficult. On the one hand, the perception of an ECA’s personality needs a long-term interaction, and even psychologist debate about how long it should take to recognize a personality. On the other hand, evaluating an ECA’s personality needs the human subjects’ collaboration and a subject may give up the evaluation if his/her personality mismatches the personality being interacted with. For example, an introverted person may give up interacting with an extroverted ECA. The initial evaluation population was proposed as being a teaching unit with 120 students but there were three reasons which made this impossible: the evaluation was to be done in the summer holiday, not all

computers had an audio output device, and the noisy Talking Head may affect the tutorial of the unit.

Additionally, according to Curtin's ethical requirement, evaluation participants are allowed to leave the study at any time without reason if they do not want to continue.

These reasons led to a reduced evaluation population that would consist of staff and post-graduate students who would use the system on the desktop computers over the summer break. The problem of the evaluation population size and makeup, along with the data analysis, will be discussed in detail in Section 7.2.

3.5 Methodology

3.5.1 Design and Implement Methodology

Action Research has been widely used for computer science design and implementation, and it was adopted as the *design and implement* methodology for this research. Micro cycles were involved in the action research cycle (see Figure 3.4(a)). Figure 3.4 (b) shows the micro cycles for the "appearance designing" module: At first, the module was planned based on the psychology requirement of personality characteristics, then the module was design and implemented within Curtin's ECA system. After that, the module was tested and finally a micro evaluation was done by the researchers, and reflections on the design were obtained.

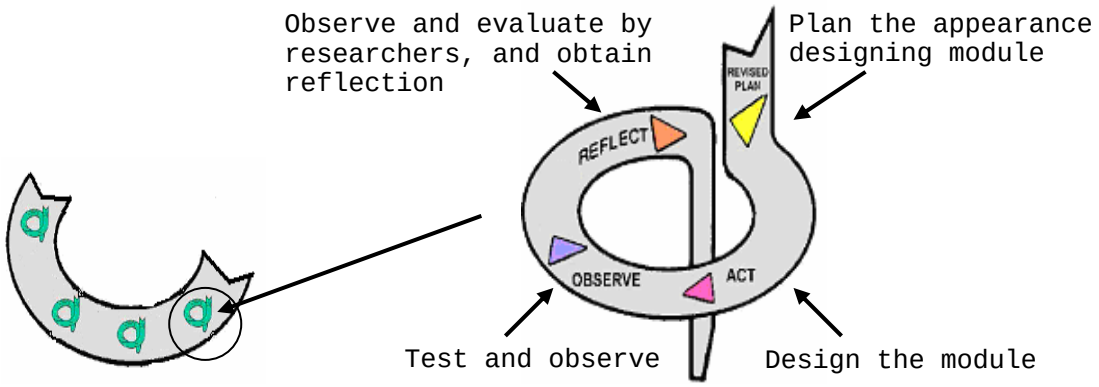


Figure 3.4: (a) Design and Implementation in the Action Research Cycle and (b) Example Minimal Action Research Cycle

Software engineering methodologies, such as Object Oriented (OO), black box test and white box test, were also adopted in the design and implementation of the computational personality model:

- The use of Java indicated an Object Oriented approach to the system design. Wilkie (1993) indicated that the most fundamental advantage of object-oriented techniques is in creating a more modular approach to the design and implementation of software systems and the biggest disadvantage is the immaturity of the techniques and tools.
- Black-box and white-box test (Beizer 1990) were used as system testing methods. Black-box testing was used to test the functionality of the system based on the system specification. For example, the testing of dialogue, animation and speech adopted the black-box test. White-box testing was used when testing the feelings, thoughts and behaviours modules in the computational model.

3.5.2 Evaluation Methodology

In order to evaluate the personality model in terms of the hypotheses, questionnaires were needed, and *evaluation* research methodologies were applied to analysing the questionnaire results. The evaluation of a personality model is not as easy as the evaluation of emotions since personality cannot be identified visually. A long-term frequent-interaction scenario - "Email Agent" - was developed as an evaluation application, and several experiments were planned. The evaluation was done by users continually interacting with the Email Agent over a two-week period, and then by answering question about the experiment. The experiment and the questionnaires were designed according to the requirements of the evaluation, which tightly related to the hypotheses. The evaluation will be detailed in Chapter 7.

A user's personality traits were surveyed in order to investigate the relationship between the user's personality, their questionnaire results, and the ECA's personality. It is important to note that standard psychological tests (Goldberg 2005) were used to evaluate the user's personality as well as the user's perception of the ECA's personality. That is, the user is evaluating the accuracy of the implemented personality model via standardised unbiased tests.

As Figure 3.5 shows, five levels, from *very inaccurate* to *very accurate*, were used to describe how accurately each statement described a user, and the user's personality characteristics can be calculated from the levels from different statements (see Chapter 6). For example, the extraversion value will increase if the user chooses "Very Accurate" for the statement "Am the life of the party" and the agreeableness value will decrease if the user chooses "Very Accurate" for "Feel little concern for others."

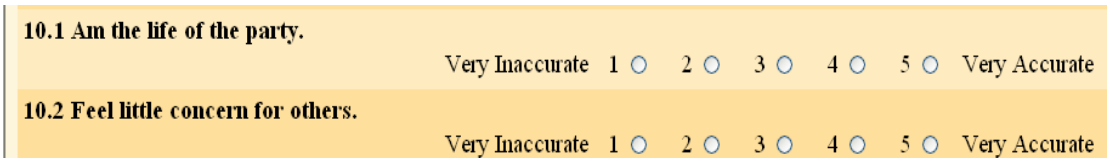


Figure 3.5: Example Questions in Questionnaire 1 (see Appendix B for the Questionnaire)

The evaluation of the personality model will analyse the users' attitude to two personality rich ECAs. It may be useful to compare a personality rich ECA to a non-personality ECA but it is not considered in this research for the following reasons:

- The population size is very low and hence cannot afford to split the groups even further.
- It has been proposed that personality rich ECA are more believable (André et al. 1999; Barker 2003; Bartneck 2002; Breese & Ball 1998; Egges & Kshirsagar 2004; Kshirsagar 2002).
- Previous studies (Stallo 2000) Shepherdson (2000) have shown that any improvement through adding personality/emotions to an ECA make the interaction more enjoyable and engaging.

It is important to understand that the evaluation methodology "mimics" what humans do in evaluating other humans. We do not ask a psychologist to determine our friend's personality. We determine it by evaluating the friend's responses to stimuli such as questions and situations. The evaluator does the same to the ECA. This study proposes that the implemented ECA personality model, based on IPIP (Goldberg 2005) criterion (see Section 4.2.2.2), will be correctly determined by the user and then used to rate the implementation correctness.

3.6 Data Collection and Storage

There are two types of data that were collected from two sources:

- *Electronic Usage Data*: users' input requests to the ECA, the responses from the ECA and the reaction of the users.
- *Questionnaire Data*: the qualitative and quantitative data from the user questionnaires.

The collected data were analysed to see whether they confirmed or denied the hypotheses and all data are included on the CD-ROM of Appendix G. The thesis and data will be retained at the Curtin library and Department of Computing for at least five years.

3.7 Significance

A personality model is a significant step forward in building a believable ECA. Research on personality models can enhance an ECA's believability since an ECA's emotion, expression and behaviour needs to be guided and controlled consistently by a personality model.

A personality model advances ECA research in that a believable ECA needs personality but few implemented personality models are currently available. Researchers have concentrated on emotion instead of personality, and created many emotional agents based on the OCC and other models. Yet personality theories such as Eysenck and Eysenck's (1963) theory and McCrae and John's (1992) Five-Factor Model do exist.

The personality model will be a strong base for further research. More complex personality models can be developed based on the personality model and many more believable personality-rich characters will be created based on these models. This research will contribute to the knowledge of ECA research in the human-computer interaction area.

The personality model will be integrated into the existing Dialogue Management system and hence future users of the DM will benefit from this personality research.

One final important aspect of ECA research is in evaluating the effect of an ECA personality on the user. This research considers how the evaluation might occur. In

many ways this is a pilot study. “How can personality be measured?”, and “How long does it take to detect personality?” were key issues to be considered.

3.8 Ethical considerations

There were members of the public involved in the experiment, and these users filled out questionnaires. Further, during their interaction with the ECA, statistics have been collected. The experimental system has stored information about a user in order to provide better service, and to allow a user to have access to only their information. The user’s name and associated information was only identifiable to the researcher and the user’s name was not important to the research. The experiment was only started after ethical approval was given by Curtin’s Human Research Ethics Committee (HREC).

3.9 Conclusion

In summary, this chapter has discussed issues that related to the RM of this research: hypotheses, research objectives, limitations, design and implementation methodologies, evaluation methodologies, data collection and storage, significance and ethical considerations. These issues have defined the research to be presented in this thesis and most of these issues will be discussed again in the following chapters. The following chapters will discuss issues in each phase of this research: personality modelling, implementation, evaluation, data analysis and conclusion.

Chapter 4

Personality Modelling

4.1 Introduction

This chapter discusses the issue of personality modelling that relates to the first two objectives of this research.

- The Psychological Modelling section (4.2) will discuss the issue of designing a specification for a personality model based on existing research in the psychology literature. Traits are discussed as the basis of the psychological personality model. Then the personality characteristics are extracted from the analysis of the psychological description. The psychological personality model is described as modules of *feelings*, *thoughts* and *behaviour* based on the characteristics.
- The Computational Personality Model section (4.3) will discuss the conversion of the psychological personality model into a computational model. The Computational Personality Model has both server-side and stand-alone modules. Modules in the psychological personality model were converted to computational modules of *feelings*, *thoughts* and *behaviours*.

4.2 Psychological Modelling

4.2.1 Traits

A real person's personality can be described by his or her family members, friends or colleagues when they answer the question: "What is he or she like?". The answers to this question may cover hundreds of adjectives such as friendly, joyful, sociable, easy-going, active etc. These adjectives indicate different characteristics of personality. Any proposed personality model will find it difficult to enable an ECA to

have all the possible personality characteristics listed as adjectives in our dictionary since most of the adjectives are abstract, and the list can be long. It is necessary to give the personality model a delimitation that covers limited characteristics for modelling and measure. This delimitation is traits.

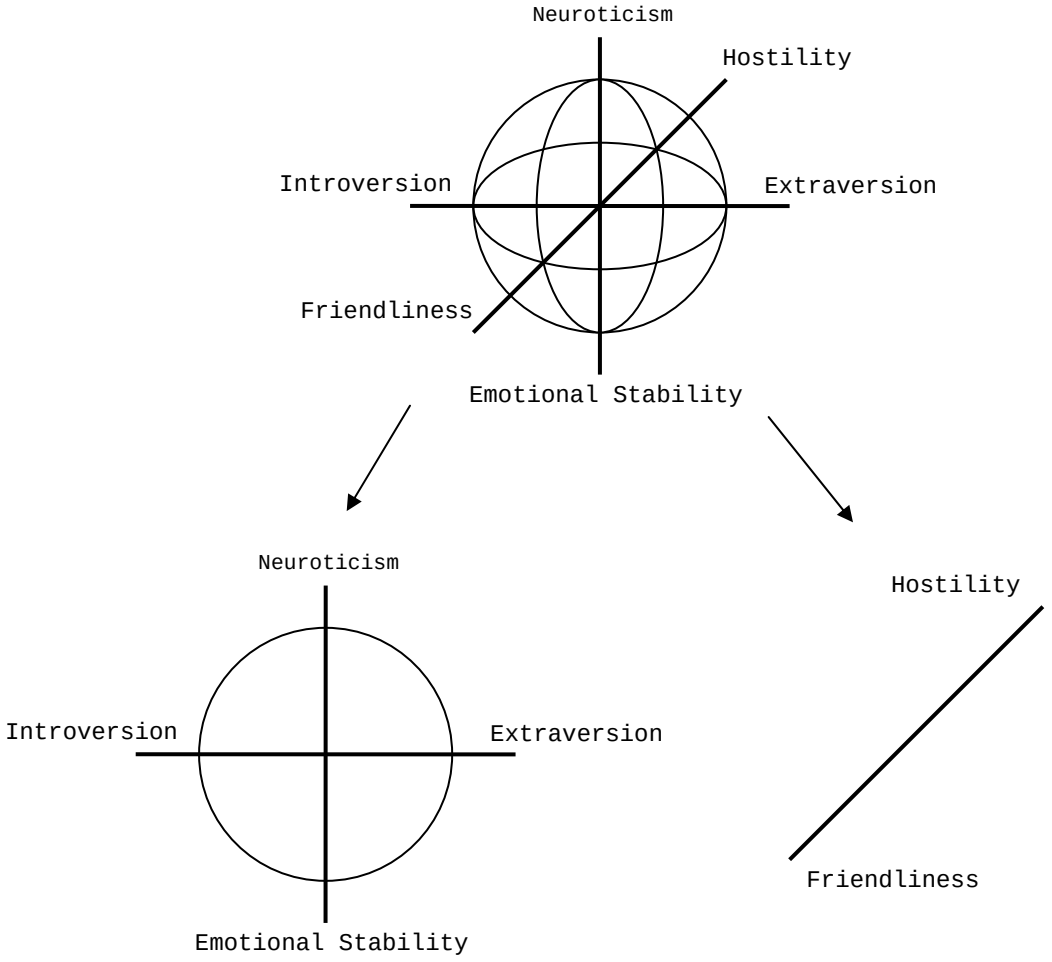


Figure 4.1: The *Friendliness, Extraversion* and *Neuroticism (FEN)* **traits of the personality model in this thesis**

Various trait theories were discussed in Chapter 2 and this research will adopt the traits in Eysenck's theory (Eysenck & Eysenck 1963) and interpersonal traits theory (Leary 1957) (see Figure 2.4). The *friendliness* trait was chosen since this is directly related to any user interaction. The trait of e*xtraversion* was chosen since it is the most relevant trait to the interaction domain of this research. The trait of *neuroticism* was adopted since it is useful for controlling an ECAs' emotions and moods (Egges & Kshirsagar 2004). The trait of *dominance* in interpersonal theory is not considered in this research since an ECA has less chance to display this kind of characteristics in the

interaction domain of "user types a question and the ECA generates an answer". The three traits, *Friendliness*, *Extraversion* and *Neuroticism (FEN)*, are used as the delimitation of the model's scope in the personality model of this thesis. These traits have also been used successfully in previous research (Kshirsagar 2002).

The evaluation of a user's perception of personality was also based on FEN traits and the characteristics involved in the traits. However, FEN traits do overlap since these traits come from two different trait theories. So the trait of *friendliness* is modelled separately while *extraversion* and *neuroticism* are modelled in a two dimensional space (see Figure 4.1). The FEN traits were defined using three dimensions in the Computational Personality Model only to simplify the system design.

4.2.2 Personality Characteristics

Concrete descriptions of the traits are needed for modelling, since traits are high level and not easily modelled as concrete characteristics of feelings, thoughts or behaviour (Pervin & John 2001). For example, an ECA that always smiles can indicate a friendly personality and an ECA that always uses long answers may indicate an extravert personality. The modelling of *smile* and *long answer* is much easier than modelling *extravert* and *friendly* directly. However, *long answers* are not enough to reflect an extravert personality, and much more concrete characteristics such as *keeps moving* and *likes to make friends* are needed to be modelled to reflect a human like extravert personality. Psychologists have described these traits in detail and the descriptions include some useful personality characteristics. Two methods were used to analyse the psychological description and to extract personality characteristics:

- Analyse the psychological description of typical personality.
- Analyse the Scales and Items of International Personality Item Pool (IPIP) (Goldberg 2005).

4.2.2.1 Analyse the psychological description of typical personality

The description of typical *extravert*, *introvert*, *neurotic* and *emotionally stable* personalities comes from Eysenck et al's (1963) personality research. For example, the typical *extravert* personality is described as follow:

"The typical extravert is sociable, likes parties, has many friends, needs to have people to talk to, and does not like reading or studying by himself. He craves excitement, takes chances, often sticks his neck out, acts on the spur of the moment, and is generally an impulsive individual. He is fond of practical jokes, always has a ready answer, and generally likes change; he is carefree, easy-going, optimistic, and likes to 'laugh and be merry.' He prefers to keep moving and doing things, tends to be aggressive and lose his temper easily: altogether his feelings are kept under tight control and he is not always a reliable person." (Eysenck & Eysenck, 1963)

Figure 4.2: The Original Description of Typical Extravert Personality (Eysenck & Eysenck 1963)

Two steps were involved in the analysis of the psychological description:

- The characteristics of an *extravert* personality were extracted from the original psychological description. The underlined words shown in Figure 4.3 are characteristics of the *extravert* personality. However, not all the characteristics can be used for modelling the *extravert* personality. These characteristics need to be revised.

"The typical extravert is *sociable*, *likes parties*, *has many friends*, *needs to have people to talk to*, and *does not like reading or studying by himself*. He *craves excitement*, *takes chances*, often *sticks his neck out*, *acts on the spur of the moment*, and is generally an *impulsive individual*. He *is fond of practical jokes*, *always has a ready answer*, and generally *likes change*; he is *carefree*, *easy-going*, *optimistic*, and likes to '*laugh and be merry*.' He prefers to *keep moving* and *doing things*, *tends to be aggressive* and *lose his temper easily*: altogether *his feelings are kept under tight control* and he is *not always a reliable person*." (Eysenck & Eysenck, 1963)

Figure 4.3: Characteristics Extraction of Extravert Description

- Revise the list of the personality characteristics based on the relationship to the ECA domain and the difficulty of modelling.

  Firstly, some characteristics, such as *sociable* and *easy-going* which are traits of other trait theories, should be removed. They can be modelled in a similar way to modelling the *extravert* personality. However, the personality model may grow unbounded in this way, and hence cannot be realised in a limited amount of time. Therefore these kinds of characteristics need to be removed.

  Secondly, some characteristics such as *places great value on ethical standards*

and *likes well-ordered mode of life* (see Figure 4.4) are too abstract to be modelled and have to be modified. The characteristics after revision are shown in Figure 4.4. Further processing of these characteristics will be done in order to develop the psychological personality model and this process will be discussed in Section 4.2.3.

"The typical extravert is sociable, *likes parties*, has many friends, *needs to have people to talk to*, and *does not like reading or studying by himself.* He craves excitement, takes chances, often sticks his neck out, acts on the spur of the moment, and is generally an impulsive individual. He *is fond of practical jokes*, *always has a ready answer*, and generally *likes change*; he is *carefree*, easy-going, *optimistic*, and likes to '*laugh and be merry*.' He prefers to *keep moving* and doing things, *tends to be aggressive* and *lose his temper easily*: altogether *his feelings are kept under tight control* and he is not always a reliable person." (Eysenck & Eysenck, 1963)

Figure 4.4: Characteristics of Extravert Description after Analysis

The analysis of the descriptions of *introvert*, *neurotic* and *emotionally stable* was done using the same methods as above and the characteristics are shown in Figure 4.5. The analysis of *friendliness* is not considered in this section and it will be discussed in Section 4.2.2.2. These characteristics are useful not only for personality modelling but also for the evaluation of the user's perception of the ECA's personality. Personality characteristics are easier to understand by an evaluator than traits since evaluators are not psychologists. For example, users, after interacting with the ECA for a long term, were asked: "Do you agree the ECA likes to talk to a lot of different people?" (see Appendix C) instead of "Do you agree the ECA is extravert?". The word "extravert" may not be understood fully by evaluators since some of the evaluators have English as their second language. The question "Do you agree the ECA likes to talk to a lot of different people?" is a standard psychological question (Goldberg 2005) and an evaluator can understand it clearly since this is a popular sentence. The ECA has no chance to talk to other people but if a user can recognize the ECA's personality correctly, the user can judge what the ECA would like to do and what it would not like to do.

“The typical introvert is a *quiet*, retiring sort of person, introspective, *fond of books rather than people*; he *is reserved and distant except to intimate friends*. He tends to plan ahead, ‘*looks before he leaps*’ and distrusts the impulse of the moment. He *does not like excitement*, takes matters of everyday with proper seriousness, and likes a well-ordered mode of life. He *keeps his feelings under close control*, *seldom behaves in an aggressive manner*, and *does not lose his temper easily*. He is reliable, somewhat *pessimistic*, and places great value on ethical standards.” (Eysenck & Eysenck, 1963)
“High neuroticism scorers are generally *anxious*, *worrying* individuals, *moody* and *frequently depressed*. He is likely to sleep badly, and to suffer from various psychosomatic disorders. He is overly emotional, *reacting too strongly to all sorts of stimuli*, and finds it *difficult to get back on an even keel after each emotionally arousing experience*. His *strong emotional reactions interfere with his proper adjustment*, making him *react in irrational sometimes rigid ways*. If the high neuroticism individual has to be described in one word, one might say that he is a *worrier*; his main characteristic is a constant preoccupation with things that might go wrong, and a *strong emotional reaction of anxiety to these thoughts*.” (Eysenck & Eysenck, 1963)
“The stable individual tends to *respond emotionally only slowly and generally weakly*, and to *return to baseline quickly* after emotional arousal; he is *calm*, even-tempered, *controlled* and *unworried*.” (Eysenck & Eysenck, 1963)

Figure 4.5: Characteristics of Introvert, Neurotic and Emotional Stable Description

4.2.2.2 Analyse the Scales and Items of International Personality Item Pool (IPIP)

The International Personality Item Pool (IPIP) is a computer-supported system that allows scientists to work with each other without regard to geographical location for the development of measures of personality and other individual differences (Goldberg 2005). The personality scales of IPIP include traits of different trait theory and each scale points to IPIP items in different psychological questionnaires (Goldberg 2005). IPIP is a freely available system and it includes descriptions of Eysenck's theory and interpersonal traits theory that the personality model is based on. These items were used in questionnaires by psychologists to measure people’s personality. As Table 4.1 shows, the items marked by “+ keyed” are positive items that indicate the increase of the friendliness value and items marked by “- keyed” indicate a decrease of the friendliness value.

Table 4.1: Items of the Friendliness Scale (Goldberg 2005)

	Items
+ keyed	Make friends easily.
	Warm up quickly to others.
	Feel comfortable around people.
	Act comfortably with others.
	Cheer people up.
– keyed	Am hard to get to know.
	Often feel uncomfortable around others.
	Avoid contacts with others.
	Am not really interested in others.
	Keep others at a distance.

IPIP items are another useful source of personality characteristics used in this study. There are three processes to obtain useful characteristics:

- Choose IPIP scales that relate to the FEN traits. Scales such as *Friendliness*, *Introversion*, *Extraversion*, *Emotional Stability* and *Neuroticism* have been chosen.
- Check IPIP items for each scale. Appendix E shows item tables of each scale.
- Move negative items to opposite scales. As Table 4.1 shows, negative items such as "Keep others at a distance." are not a characteristic of a friendly person, and can be moved as a characteristic of a hostile personality.
- Cut items out if they are too abstract, too difficult for modelling in the limited time frame, or useless for the ECA domain. Some obvious items such as "Get upset easily" and "Often in a bad mood" were adopted for modelling after the above processes.

IPIP items were also used to measure the personality traits of the evaluators before they took part in the evaluation - see Chapter 7 and Appendix B for detail.

4.2.3 Psychological personality model

The previous section discussed the extraction of personality characteristics of FEN traits from the psychological literature. A psychological personality model needs to be designed in order to model these characteristics. *Feelings*, *Thoughts* and *Behaviour* are the three parts of the model, and personality characteristics need to be classified into these parts (see Figure 4.6). These three parts come from the broad definition of personality as patterns of thinking, feeling and acting (Pervin & John 2001).

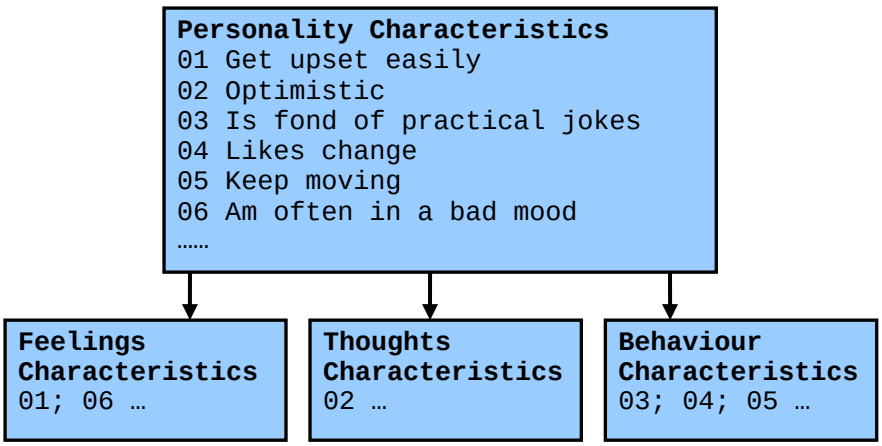


Figure 4.6: Classification of Personality Characteristics

Figure 4.7 shows the psychological personality model of this research. Modules such as "Idle behaviour" and "Change appearance" are computational modules of the personality model, and result from the analysis of the personality characteristics in Figure 4.6. The psychological personality model links psychological personality characteristics to the computational personality model.

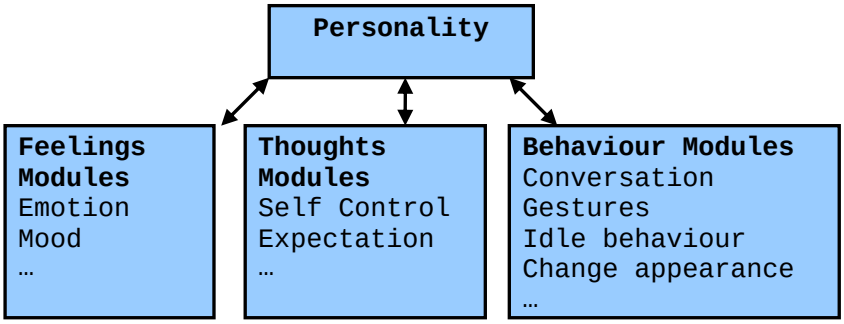


Figure 4.7: Psychological personality model

Feelings Module: Personality characteristics such as "Am often in a bad mood", "Get upset easily" and "Get angry easily" belong to *feelings*. They are either related to emotion or related to mood and therefore emotions and moods are considered as the main elements of feelings. Feelings also include some processes of emotions and moods such as appraisal, expression and decay (see Section 4.3.1).

Thoughts Module: Some abstract personality characteristics descriptions, such as "places great value on ethical standards" and "likes well-ordered mode of life", have been revised in Section 4.2.2. Some abstract characteristics that are related to thoughts are also possible to model. For example, "optimistic" and "pessimistic" can be

modelled by the expectation of an agent. An optimistic agent has good expectations while a pessimistic agent may have bad expectations. Modelling the expectation of an agent is easier than modelling optimism since expectation has been widely discussed in emotional research such as Ortony's OCC model (Ortony, Clore & Collins 1988) and Bui's plan-based emotion model (Bui et al. 2002).

Behaviour Module: Characteristics such as "Is fond of practical jokes", "Likes change" and "Keep moving" belong to *behaviour*. These characteristics normally are concrete and can be modelling using ECA techniques such as conversational and non-verbal, non-conversational behaviour, gestures, facial animation and idle behaviour.

4.3 Computational Personality Model

Curtin's ECA system is client-server based, with the DM, KB and TTES on the server side and the FAE and User Interface on the client side. Curtin's previous personality models did not fully integrate into the TH – it was always an add-on either to the server or to the client. The current research has concentrated on developing an integrated framework that can be used by both the client and the server. Figure 4.8 is a schematic of the new affective computational personality model.

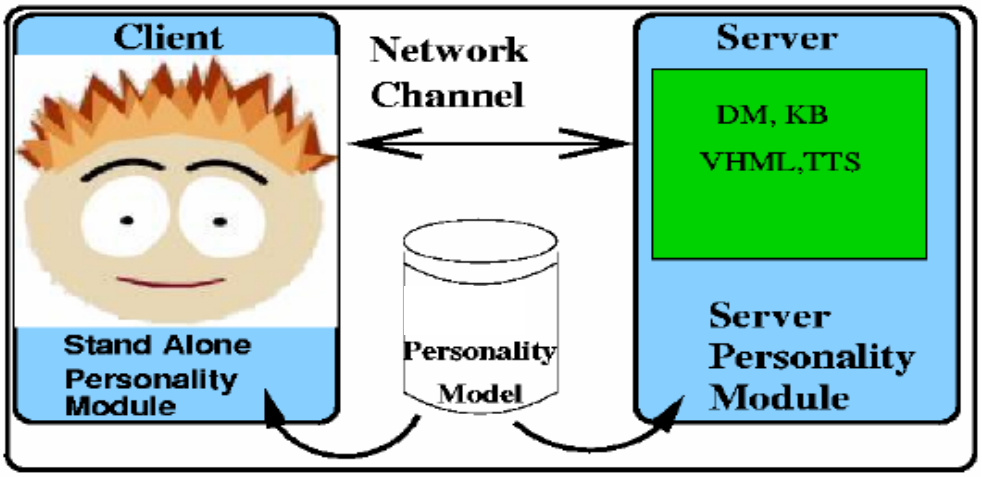


Figure 4.8: The Computational Personality Model

The Computational Personality Model is implemented as Object Oriented code that is loaded and executed at runtime on both the server and client. The network channel allows both the system and the user personality models to be used interchangeably, with the relevant code being either sent from the server to the client or vice versa. The ability to create and load client side user personality models allows for transparent

research by disparate groups. The Computational Personality Model can "remember" user interactions through the use of per-user files that contain learnt preferences about how to interact with each user. The user interaction can change over time as the ECA personality changes the per-user file preferences.

Server side personality modules can control the *feelings*, *thoughts* and *behaviours* of an ECA when the client can access the server. However, if no connection is available to the server, a stand-alone personality model will work. Server side personality modules include modules of *feelings*, *thoughts* and *behaviour,* while the stand-alone personality model only has an active response module, an idle behaviour module and an appearance-designing module (see Figure 4.9). These modules will be discussed in following sections.

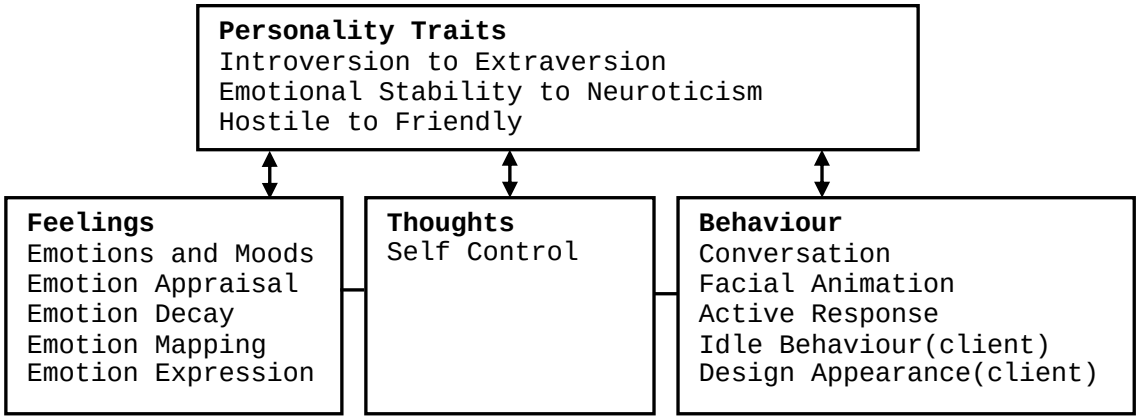


Figure 4.9: The Components of the Computational Personality Model

4.3.1 Feelings

The psychology literature (Allport 1937; Eysenck & Eysenck 1963; Hewstone, Fincham & Foster 2005; Myers 2004) suggested that one of the functionalities of a personality is that it guides a human's feelings. Feelings include affective state of consciousness, physical sensation, awareness and impression. Although feelings and emotions are sometimes interchangeable, feelings are the more general and neutral (Farlex 2005). The computational personality model used emotions and moods as two main elements of feelings for the following three reasons:

- Emotions and moods were widely researched by computer science researchers (Breese & Ball 1998; Reilly & Bates 1992) and more suitable for modelling than feelings.

- Emotions and moods are required by the psychological personality model, since the personality characteristics in the category of feelings either related to emotions or related to moods.
- Emotions and moods are requirements for a believable ECA (Kshirsagar 2002; Ruttkay, Dormann & Noot 2002).

Ekman's (1972) six basic emotions were adopted as emotion states for two reasons: Firstly, the physical appearance of Curtin's ECA only covers the head, and facial animation is the main way to express emotions. So Ekman's emotions were chosen since they are ideal for facial expressions. Secondly, the 22 emotion states in the OCC model are too many for the domain of interaction.

The computational definition of emotions and moods is shown in Figure 4.10. Emotions were defined as states and intensities. As Figure 4.10 shows, each state is followed by an intensity value that ranges from 0 to 100. Moods were defined as three states: bad mood, neutral and good mood.

```
public String Emotions[][] =
{     //emotion state, emotion intensity
      {"angry",          "10"  },
      {"happy",          "20"  },
      {"disgusted",      "0"   },
      {"afraid",         "0"   },
      {"suprised",       "0"   },
      {"sad",            "0"   }
};

public int Mood;
static final int BADMOOD = -1;
static final int NEUTRALMOOD = 0;
static final int GOODMOOD = 1;
```

Figure 4.10: Emotions and Moods

4.3.1.1 Emotion Appraisal

The appraisal of an ECA's emotions and moods is based on the emotion processes suggested by Bartneck (2002) (see Section 2.2.3). The following steps were involved in the appraisal of emotions in the personality model:

Classifications: Find what emotional states are affected based on the ECA's personality and the user's question. In the Mentor system (Marriott 2005), each

question from the user will be matched to a Domain topic. For example, “How are you?” is matched by a topic called greetingTopic and “Bastard” is matched by the swearTopic. Emotional states can be found by analysis of the topic that matches the user’s question.

- Some topics always result in a specific emotional state regardless of the ECA’s personality. For example, “greeting” or “polite” topics make ECA happy and the “swear” topic always makes the ECA angry. Questions that belong to these kinds of topics match the fixed emotional states. So “How are you?” or “How do you do?” matches the state of “happy”.
- Feelings of some topics are related to the personality of an ECA. For example, questions like “2001 + 1788 =” belongs to the “arithmetic” topic and this topic may make a friendly extravert ECA happy whilst it may make a hostile introvert ECA angry. Questions that belong to this kind topic match the emotional states according to ECA’s personality traits. Although an extremely hostile person may be angry with any conversation, it will be useless modelling a hostile personality that is angry with every topic. The hostile personality is modelled as not always being angry but upset easily. For example a hostile ECA is unlikely to become angry with “how are you?”.
- Feelings of some topics are related to the content of the questions. A statement such as like “I like Alice” is processed by the “I like” topic and one ECA may be happy with “I like Alice” since he likes Alice, and yet cause anger with “I like David” if the ECA does not like David. Similarly for statements such as “I hate William” or “I find William boring”. Questions in these kinds of topics match emotional states according to the content of the user’s questions and the ECA’s preferences. The appraisal is necessary since the response needs to be tagged by the current emotion “<angry>No, I hate Alice<blink/></angry>” in order to achieve an emotional speech with animation.

Additionally, the mood of an ECA may change when a specific emotion is caused by a specific event such as a neurotic ECA changing to a bad mood when it feels very angry by the user’s rude words. Users’ questions do not cause the update of emotions and moods if they do not match any emotional state.

Quantification: Obtain the increments of intensities of the affected emotional state and mood by reading the ECA's knowledge and then use the increments to update the ECA's current emotional state and mood. The emotion and mood increments are recorded as ECA knowledge in the personality model. It is assigned in this simple manner because of the scope/duration of the one-year research. Translation from events to emotions' increments or decrements, emotion decay and emotion mapping will be described in this section, and in Section 4.3.1.2.

```
//topics, emotional states, emotion increment, mood increment
"greeting",  "happy",              "25",             "0"
```

The value of "25" in the table entry above represents the emotion increment – a value from zero to one hundred. Zero means no increase while one hundred means extreme increment.

The value "0" above is the mood increment which can be "-1", "0" or "1". As Figure 4.11 shows, the mood can be updated by adding the increment to the current mood (line 1). Lines 2 and 3 ensure that the mood only has three values -1, 0 or 1, which indicates a *bad mood*, a *neutral mood* or a *good mood*. The actual increment of emotion intensity is not simply equal to the emotion increment that was recorded in the personality model. For example, a happy event may make a happy person a bit happier while it may make a person with neutral emotion much happier. So the actual increment is affected by both the current emotion, the emotion increment, and the function $F(E_{t-1}, E_i)$ as shown in line 5: update the current emotion by using the $F(E_{t-1}, E_i)$ increment function as opposed to adding E_i directly (see Figure 4.12).

$$M_t = \begin{cases} M_{t-1} + M_i & \text{line 1} \\ -1 \quad \text{if } M_{t-1} = -1 \text{ and } M_i = -1 & \text{line 2} \\ 1 \quad \text{if } M_{t-1} = 1 \text{ and } M_i = 1 & \text{line 3} \end{cases}$$

$$E_t = E_{t-1} + F(E_{t-1}, E_i) \quad \text{line 4}$$

$$F(E_{t-1}, E_i) = (100 - E_{t-1}) * (E_i / 100) \quad \text{line 5}$$

E_t Emotion after update — M_t Mood after update
E_{t-1} Current emotion — M_{t-1} Current mood
E_i Emotion increment — M_i Mood increment
F() Increment Function

Figure 4.11: Update Current Emotion and Mood

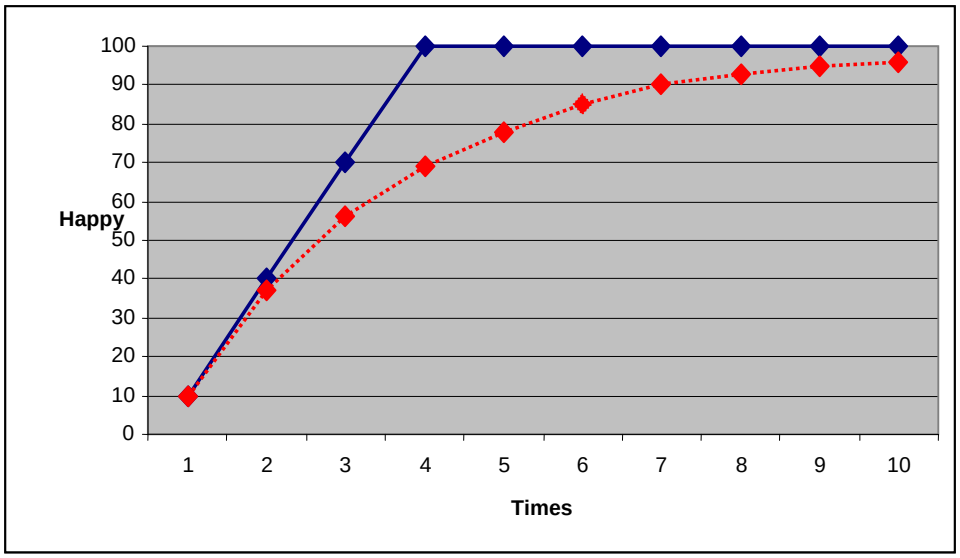


Figure 4.12: $E_t = E_{t-1} + F(E_{t-1}, E_i)$ (red) vs. $E_t = E_{t-1} + E_i$ (blue)

For example, suppose the user inputs some greeting words several times. This ECA's intensity for "happy" is updated as the red dashed line show. The first time, the ECA's "happy" intensity increased from 10 to 37 while at the last time it only increased from 95 to 97. The solid blue line shows the increase when the fixed value E_i is used as an increment. The intensity is increased the same each time until it is equal to 100. The value 100 means that the ECA feels extreme happy.

It could be assume that the red dashed line is more human like than the solid blue line but this is difficult to prove. For example, if a user keeps repeating the same question, the ECA's relevant emotion intensity should not become higher and higher, or potentially abusive or chaotic. This type of interaction was not considered in the implemented personality model.

The function of updating the intensity due to an event is not enough to accurately model the changes of emotion. The decay of emotions should also be considered.

4.3.1.2 Emotion Decay, Mapping and Expression

Emotion and mood decay: The decay of emotion depends on both personality traits and the current emotion. Firstly, the emotional states of an emotionally stable ECA will decay more than a neurotic ECA. Secondly, the emotional states with higher

intensity decay more than that with lower intensity. The decay of mood is only dependant on the personality. An ECA in a bad mood will return to neutral quickly if it has the personality characteristic “return to baseline quickly” (Eysenck & Eysenck 1963) while it stays in a bad mood for a long time if its personality has the characteristic of “often in a bad mood” (Goldberg 2005).

Emotion Mapping: Emotions need to be mapped to one emotion since Curtin’s ECAs cannot express blending emotions (see Section 2.2.3 for emotion blending). Emotions shown in Figure 4.13 can be mapped to happy since the intensity of happy is highest. As Figure 4.13 shows, the researcher assumes that the former emotion has a higher priority than the later emotion if two emotions have the same intensity. That is, if both angry and happy have the intensity of “85”, the ECA will express angry instead of “happy”.

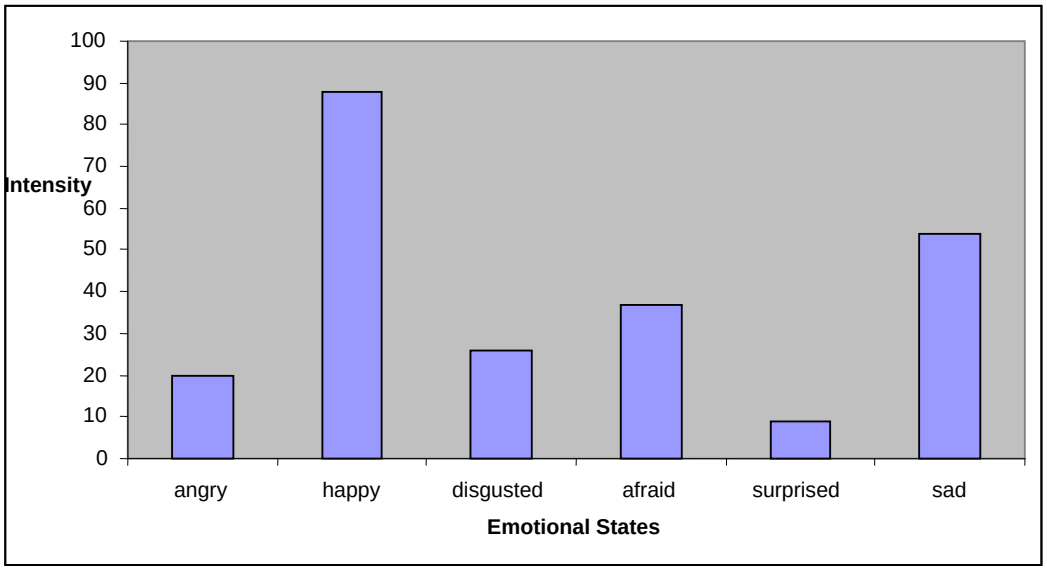


Figure 4.13: Example Emotion States and Their Intensities

Emotion Expression: An ECA can decide whether or how to express its emotion, and the expression can be via different behaviours such as conversation behaviour and/or facial animation (see Section 4.3.3).

4.3.2 Thoughts

Thoughts are more abstract than emotions and moods, and also more difficult to model. In a psychological personality model, thoughts include some concrete personality characteristics such as optimism, pessimism, expectation and self-control.

However, the *thoughts* module of the computational personality model does not implement some of these characteristics either because they are not very useful for the interaction domain, or because they are not easily modelled in the time frame of this research.

The thoughts module has two parts: self-control and conversation style. Self-control in the thoughts module can guide the expression of emotions such as whether or not to express the emotion and how to express it. For example, if a highly self-controlled ECA has a *disgusted* emotion, it may speak in a neutral way instead of using a *disgusted* speech. Its conversation may differ from ECAs with low self-control. The conversation style is decided by the ECA's thoughts. For example, an extravert ECA may like to give a long answer whilst a hostile ECA may tend to give a hostile response.

4.3.3 Behaviours

This section discusses the behaviour modules of the computational personality model. Emotional conversational behaviours such as the dialog between the ECA and the user, and the emotional speech facilities, were adopted as modules for the server side personality model. The server side personality model also involved non-verbal behaviours such as facial animation and gestures that were considered to reinforce the conversation of ECAs. Stand alone behaviour modules such as the active response module, idle behaviour module and appearance designing module were developed not only for the server side personality model but also for the client side stand alone personality model. The requirements of the different modules came from the analysis of characteristics in the psychological personality model. For example the active response module is from the characteristic "Is fond of practical jokes".

4.3.3.1 Emotional Conversational and Non-verbal Behaviours

The interaction domain of this research is: a user enters a question and an ECA responds with voice and actions according to its personality. According to this domain, the conversational behaviours can be divided into two steps:

The first step is the dialog behaviour: an ECA should know the meaning of the user's input and decide what it wants to respond to the user, and in what style. The Mentor

system (see Section 2.3.2.4) is a multi-modal dialogue manager that can match a user's questions to different topics and generate the answer from these topics. The ECA's dialog behaviour can be done by Mentor system without considering personality. That means that the content and style of the ECA's responses are random even though these responses may reflect different personality. For example, "Hi <first_name/>, what can I do for you" and "Yes Boss" reflect different personality responses as a greeting. The dialog behaviour module of the personality model decides the response and conversational style according to personality traits and feelings of an ECA, based on the processes of the Mentor system. Four conversational styles were modelled in the dialog module: long and friendly, short and hostile, long and hostile, short and friendly. These styles reflect not only personality traits but also emotional states. For example, the hostile style response comes from either a hostile ECA or an ECA currently feeling angry.

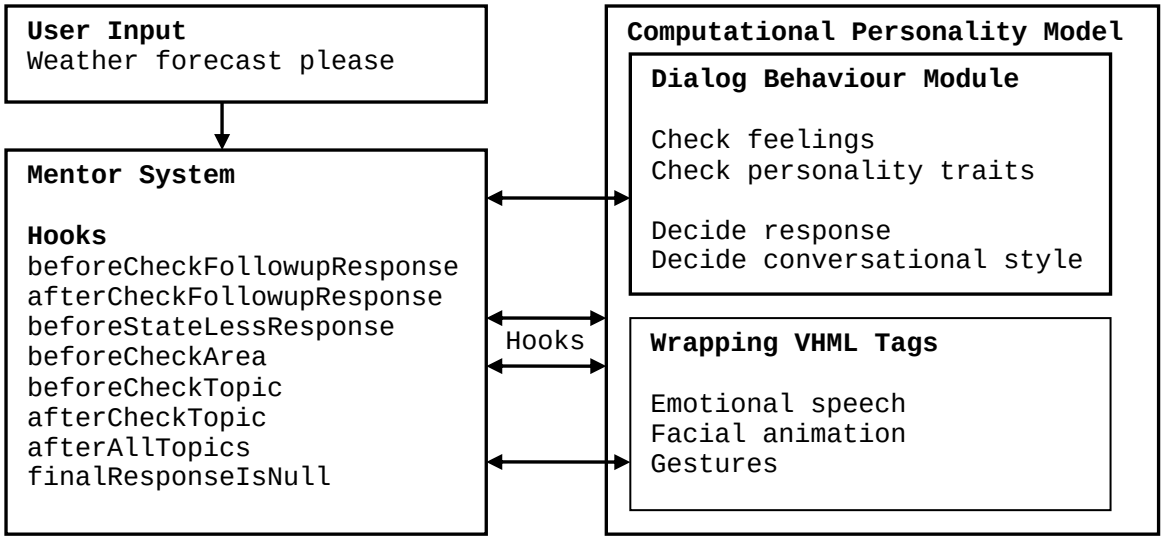


Figure 4.14: Dialog Behaviour Module and Mentor System

As Figure 4.14.shows, if the user input is "Weather forecast please" then the Mentor system will receive it and find that it matches the patterns of the weather topic. That topic may respond either with the actual weather forecast or one of the answers in Figure 4.15. Suppose that the Mentor system chooses the response "Tonight's forecast: dark. Continued dark, with scattered light around sunrise." The dialog behaviour module gets the response and the topic information through the hooks of the Mentor system. These hooks in the Mentor system were added as part of this research to provide support for any implemented Personality modules. These hooks will invoke methods in any loaded server side personality module, at various stages in

the dialogue management process. The personality module can choose to either leave the passed in information unchanged (information such as the response text), or it can change it to suit its requirements. In this case, the personality module checks the ECA's feelings and personality traits (see Figure 4.14) and the response may change to "Shut up, no weather today." if the ECA is currently angry. Or, if the ECA is an introvert, it may say "Sorry <first_name/>, I don't get out much."

```
"Tonight's forecast: dark. Continued dark, with scattered light around
sunrise.",
"Frost in susceptible areas.",
"Most parts of the country will have weather today.  Except Tasmania.",
"We'll either have sunny skies tomorrow or a plague of frogs. Not both.",
"I don't know, I haven't been outside, today.",
"No weather today. Sorry.",
"Ask me again <first_name/> and I may check the Web",
"There is a sheep wether alert!",
"Sorry <first_name/>, I don't get out much.",
"The system will be going down because of slight patches.",
"I am not interested in the weather.",
```

Figure 4.15: Responses of Weather Topic in Mentor System (Marriott 2005)

The next step in the processing of the response is that the ECA can speak the response using an emotional voice. A believable ECA should have emotional speech: saying "Shut up, no weather today." in an angry way, whilst saying "Hi <first_name/>, good afternoon." in a happy way. In Curtin's ECA system, the text to emotional speech is controlled by VHML. In this step the personality model wraps appropriate emotion tags around the plain text response and returns the VHML format response back to the Mentor system from one of the many hook routines. For example "Shut up, no weather today." is wrapped as "<angry> Shut up, no weather today. </angry>". The VHML format response will be sent to the client side personality model and finally sent to the VHML server, where it is converted into speech, and emotional gestures and expressions for the ECA. The emotional speech is outputted after the VHML server receives the VHML format response. See Section 5.2 for the relationship between the Mentor System, the VHML server and the personality model.

Non-verbal behaviours such as facial animation and gestures are also controlled by VHML. Tags such as <blink/> <smile/> <look-left> </look-left> (see Figure 4.16) indicate the blink, smile, and look left animations. These tags, added by the computational personality model based on the ECA's feelings, thoughts and personality traits, can be wrapped around the response. Figure 4.16 shows a

comparison of two ECAs interactions with one user. The responses are VHML formatted and each response is obtained through the Dialog Behaviour Module and VHML wrapping modules.

```
User: hi
ECA1: <happy> Yes boss <smile/> </happy>
ECA1: <happy> New mail from <blink/> Ivan Jackson </happy>
User: you bastard
ECA1: <angry> Same to you! </angry>
ECA1: (new mail coming) <look-left> Who cares! </look-left>
```

```
User: hi
ECA2: <happy>You disappeared for a long time!<blink/></happy>
ECA2: <happy><smile/> You have a new mail from your friend
Ivan. </happy>
User: you bastard
ECA2: <angry> You have to improve your language before
improve your learning. </angry>
ECA2: (new mail coming) <angry> You have a new mail
from Tom Lee. </angry>
```

Figure 4.16: Interaction between User and Two ECAs

4.3.3.2 Stand Alone Behaviour Modules

Idle Behaviour

It would be unnatural for the ECA to remain motionless and unblinking every time the server stops sending data. This is why the idle behaviour is needed to provide idle animation such as eye blinks, eye movement and head rolls. An idle behaviour for Curtin's ECA's was implemented in the MetaFace system (Beard 2004) using a parameterised approach. The idle behaviour module of the personality model is to control the idle time animation of MetaFace system according to the personality traits (see Figure 4.17). An extravert ECA tends to "keep moving" (Eysenck & Eysenck 1963) and it has more idle behaviours than an introvert ECA. Idle behaviour will be generated either when the client cannot access the server, the server stops sending data, or if there is a lull between interactions.

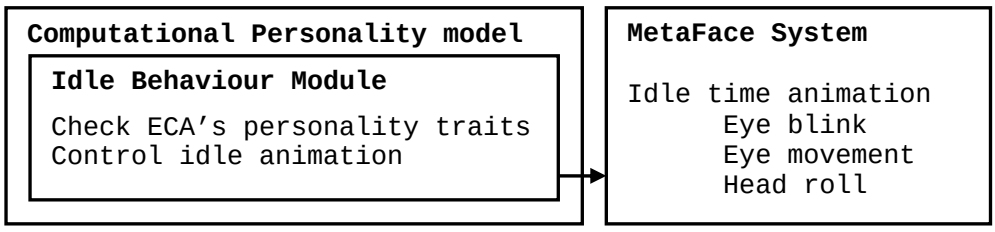


Figure 4.17: Idle Behaviour Module and MetaFace System

Active Response

Personality characteristics such as “needs to have people to talk to” and “is fond of practical jokes” indicate that some ECAs may not always keep silent when the user does not input questions. They may try to talk to the user and this behaviour is called an active response behaviour. Figure 4.18 shows an example of the active response. The generation of an active response is controlled by the personality traits of an ECA. An ECA who “needs to have people to talk to” will generate active responses more often than a quiet ECA.

```
ECA says to user:
Do you know this: The first can opener was invented in 1858 by
American Ezra Warnet. The well-known wheel-style opener was
invented in 1925.  Beer in a can was launched in 1935. The easy-
open can lid was invented by Ermal Cleon Fraze in 1959.
```

Figure 4.18: Active Response Example

Design Appearance: Self-made ECAs

Some characteristics of an ECA’s appearance, such as gender and head shape, cannot change during interaction otherwise the user will recognize it as a different ECA. However, an ECA’s appearance can be designed by itself just like humans do. An ECA should be able to design its hairstyle, lip colour, beard style or even choose to wear glasses and jewellery so as to reflect its personality. Personality often dictates a person’s appearance. Psychologists (Eagly et al. 1991; Patzer 1985; Popkins 1998) found that physical appearance is an important factor in the development of personality. Because people form opinions by what they see in a person physically, they respond to that person accordingly (Popkins 1998).

The *appearance-designing* module in Figure 4.19 is based on the reciprocal determinism of social-cognitive theories (see Section 2.2.2.4). ECAs equipped with appearance designing behaviour can be called self-made ECAs. The scope of this work is time limited, and hence minimal facial self-design was allowed to be consistent with personality. The facial self-design is minimised in order to minimise the impact of the face on the study. Only hair and lip colour designing were implemented in the *appearance-designing* module of the personality model. An extrovert ECA likes to change its appearance and tends to make itself look *cool*, while an introvert ECA does not change so much. Figure 4.20 shows the appearance over time of an extrovert ECA where the ECA itself changes its own hair and lip colour.

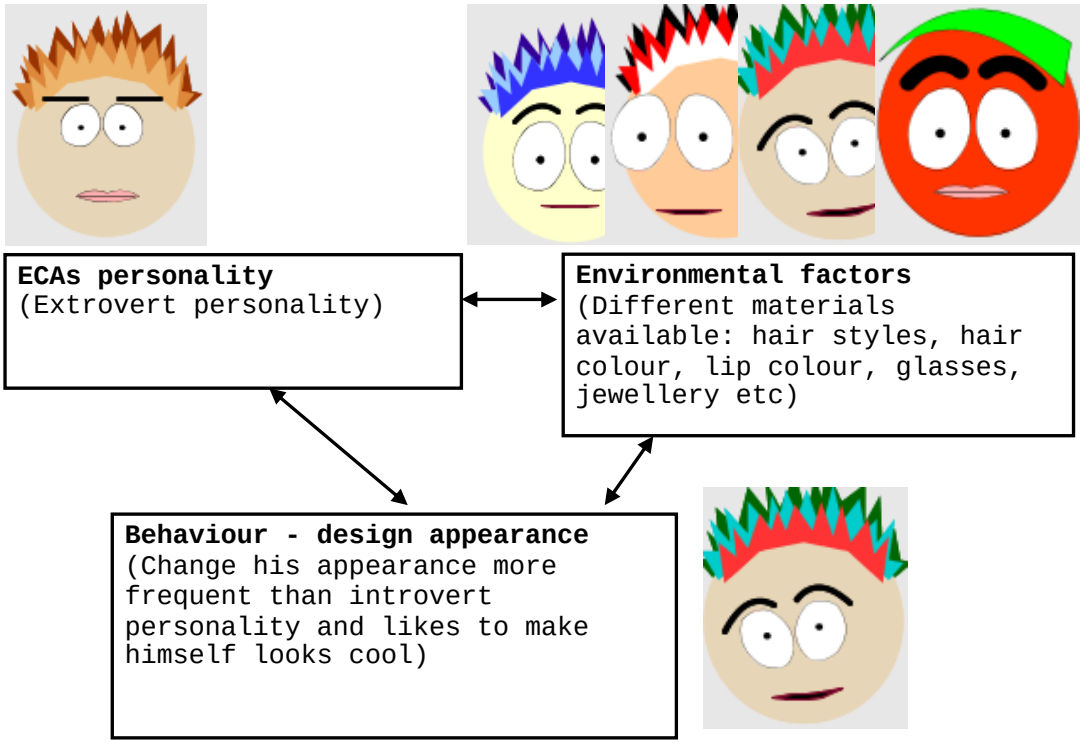


Figure 4.19: Appearance Designing Module and Self-made ECAs

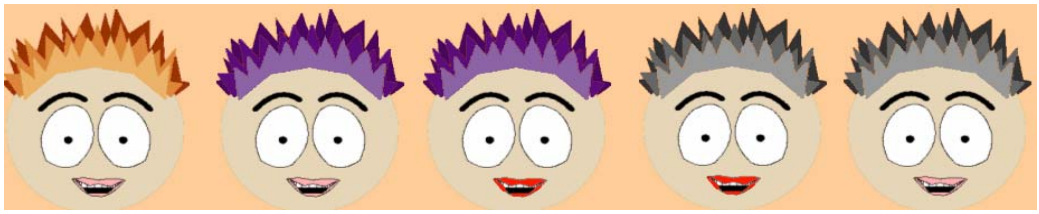


Figure 4.20: Appearance Designing Behaviour of an Extrovert ECA

4.4 Conclusion

This chapter has discussed the issues of both psychological personality modelling and a computational personality model, which relate to the two phases of this research. Section 4.2 presented the first phase of this research: design a specification for a personality model based on existing research in the psychology literature. FEN traits were described, then two methods of personality characteristics extraction were discussed, and finally the psychological personality model was introduced. Section 4.3 presented the second phase of this research: the conversion from the psychological personality model into a computational model. Computational modules of feelings, thoughts and behaviour in the computational personality model were discussed in detail.

Chapter 5
Implementation

5.1 Introduction

The previous chapter presented the Computational Personality Model. This chapter is concerned with the third objective of this research: implementation of this personality model within Curtin's ECA system. To evaluate the hypotheses of this research, an experimental application called Email Agent was developed. The Email Agent fully implemented the personality model and it ran within Curtin's ECA system. Firstly, the system architecture will be described. This includes the Mentor and VHML servers, the server and client side personality model, and the user GUI. Then the experimental application is introduced, followed by a discussion of user interface and mail monitor modules.

5.2 System Architecture

Figure 5.1 shows the overall system architecture of the implemented personality model with the grey areas representing this thesis's contribution. The system has four main components:

- **Mentor System Server** handles the dialogue management between the user and an ECA (see Sections 2.3.2.4 and 4.3.3.1). This system was already extensible but was modified to add the necessary hooks for the server side personality to be involved when the dialogue manager parses the user input and produces a response (see Figure 4.14). These hooks were used by the personality model to change the response according to the ECA's personality, and to wrap it with emotional behaviour VHML tags. The improved mentor system now has a default per-user non-interfering personality that can be over-ridden by one loaded by a user when they connect to the system.
- **VHML Server** receives the VHML format files and generates appropriate MPEG-4 facial animation and emotional speech. The interface to the VHML server is through

the MetaFace framework (Beard & Reid 2002) which manages the 2D ECAs and controls the MPEG-4 animation and idle behaviour of the ECAs.

- **Server Side Personality Model** controls the ECA's feelings, thoughts, conversational behaviours and non-verbal behaviours (see Sections 4.3.1, 4.3.2 and 4.3.3.1). The server side personality model is implemented as a standard Mentor topic, plus a personality and a personality data module. These modules (Java classes) form the base class for all future personality classes for the DM system.
- **Client Side Personality Model** controls the ECA's stand alone behaviours such as idle behaviour, appearance designing behaviour and active response (see Section 4.3.3.2).

Modules shown in grey in Figure 5.1 were developed in this research, while VHML server and Mentor system are existing systems. The UML deployment diagrams of the existing systems and the new Email Agent system are shown in Appendix F. The last deployment diagram in the Appendix F shows the designed architecture of the system, with previous work being greyed out.

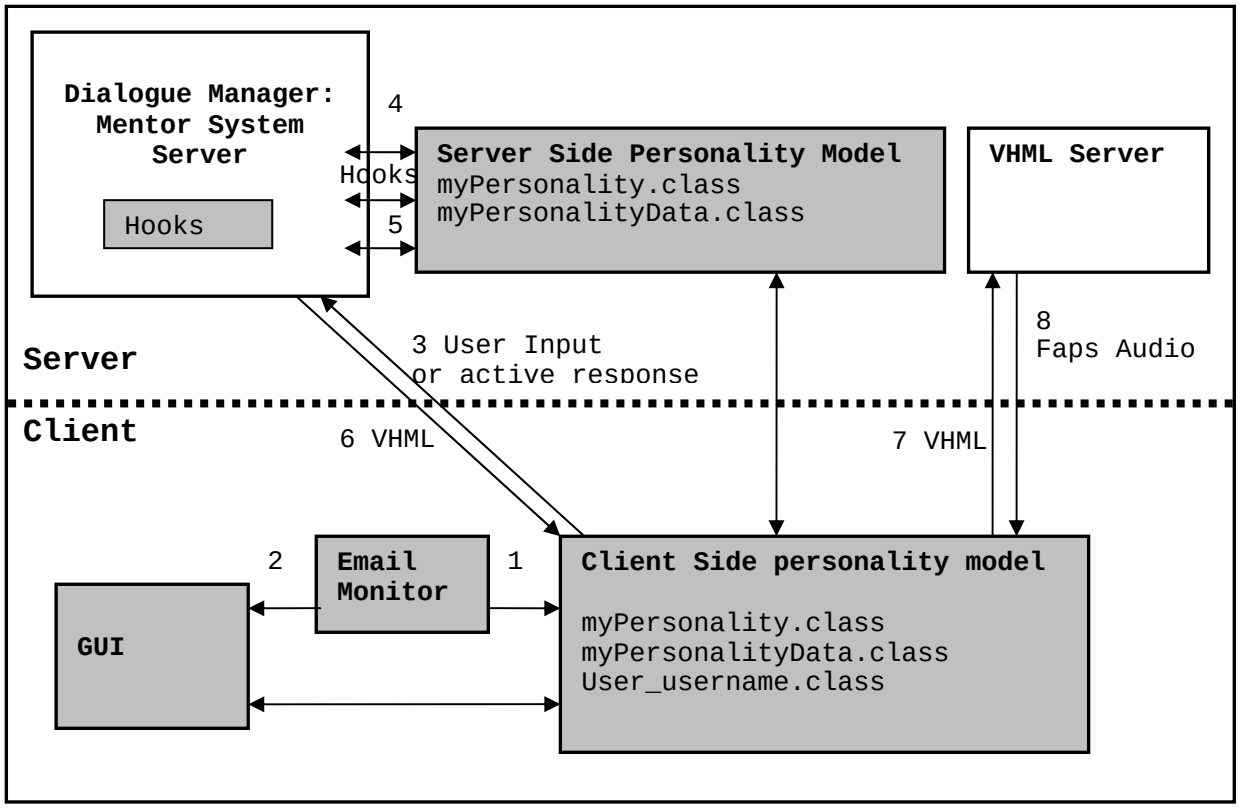


Figure 5.1: System Architecture

The architecture, except for the email monitor, is for any ECA application involving personality. The Graphical User Interface (GUI) and email monitor are components of the experimental application called Email Agent and will be discussed in Sections 5.3, 5.4 and 5.5. The numbers in Figure 5.1 indicate the data flow of the system and will be discussed in Section 5.3.

5.3 Experiment Application: Email Agent

Email Agent is an experimental application using the computational personality model and is used to evaluate the personality model. The main function of the Email Agent is to notify the user when the user receives a new email. The Email Agent can interact with the user in different ways based on its personality characteristics and its implemented modules of feelings, thoughts and behaviours, which were discussed in the last chapter. The Email Agent application can access the Mentor System and that enables it to process user's questions such as "weather forecast", "Who is Bill Clinton?" etc. ECAs with different personalities will respond differently even for the same question. The Email Agent is developed entirely using Java and it can work on Windows, Linux and Macintosh machines.

The Email Agent system includes all modules shown in the system architecture in Figure 5.1.

Personality is a long-term characteristic, so the evaluating application interaction has to be long term. Further, the user has to interact continuously with the system. Adding Knowledge Domain topics such as weather, etc allows for a general conversation with the agent to increase this interaction.

An example data flow when using Email Agent is in Figure 5.2. The numbers in Figure 5.2 match the numbered steps in Figure 5.1. The first two steps are done by the Email Monitor module that was developed from the Java Mail API (see Section 5.5). As Figure 5.1 and 5.2 shows, when it receives a new email, the Email Monitor module will send the "From", "Subject" and "Content" strings to the client side personality model and the GUI. The client side personality model extracts the sender

name and sends the “From” and “Subject” strings to the server. The personality model topic matches this formatted information and records it. As Figure 5.2 (3) shows, the sender name “Adam He Xiao” is extracted from the email “From” header “Adam He Xiao <12759585@cs.curtin.edu.au>” and “From: Adam He Xiao Subject: Test Email” is sent to the server. The server side personality model receives the “From” and “Subject” formatted input and knows that this is a new email request instead of a user’s question (see Figure 5.2 step 4).

```
1. From: Adam He Xiao <12759585@cs.curtin.edu.au> Subject: Test Email
2. Email Content: How are you?
3. From: Adam He Xiao Subject: Test Email
4. Topic: New Mail From: Adam He Xiao Subject: Test Email
5. <p><happy>You got a mail from your friend <blink/>
   Adam.</happy></p>
6. <p><happy>You got a mail from your friend <blink/>
   Adam.</happy></p>
7. <p><happy>You got a mail from your friend <blink/>
   Adam.</happy></p>
8. Facial animation and emotional speech
```

Figure 5.2: Data Flow in Email Agent

Then the server side personality model generates the plain text response “You have a mail from your friend Adam.” and adds the VHML tags “ ”, “<happy>” and “<blink/>” to generate the string “ <happy> You have a mail from your friend <blink/> Adam. </happy> ”. The VHML format response will be returned from the Mentor system and sent back to the client (see Figure 5.1 step 5 and 6).

After that, the VHML response is sent by the client to the VHML Server for conversion into speech and MPEG-4 facial animation. Finally, the client GUI receives back the MPEG-4 facial animation and emotional speech, and renders them.

5.4 User Interface

Figure 5.3 shows the user interface of the Email Agent GUI. The GUI was developed using the Java Swing Toolkit, which takes advantage of its portability and object-oriented features. The Email Agent GUI includes the following elements (see Figure 5.3):

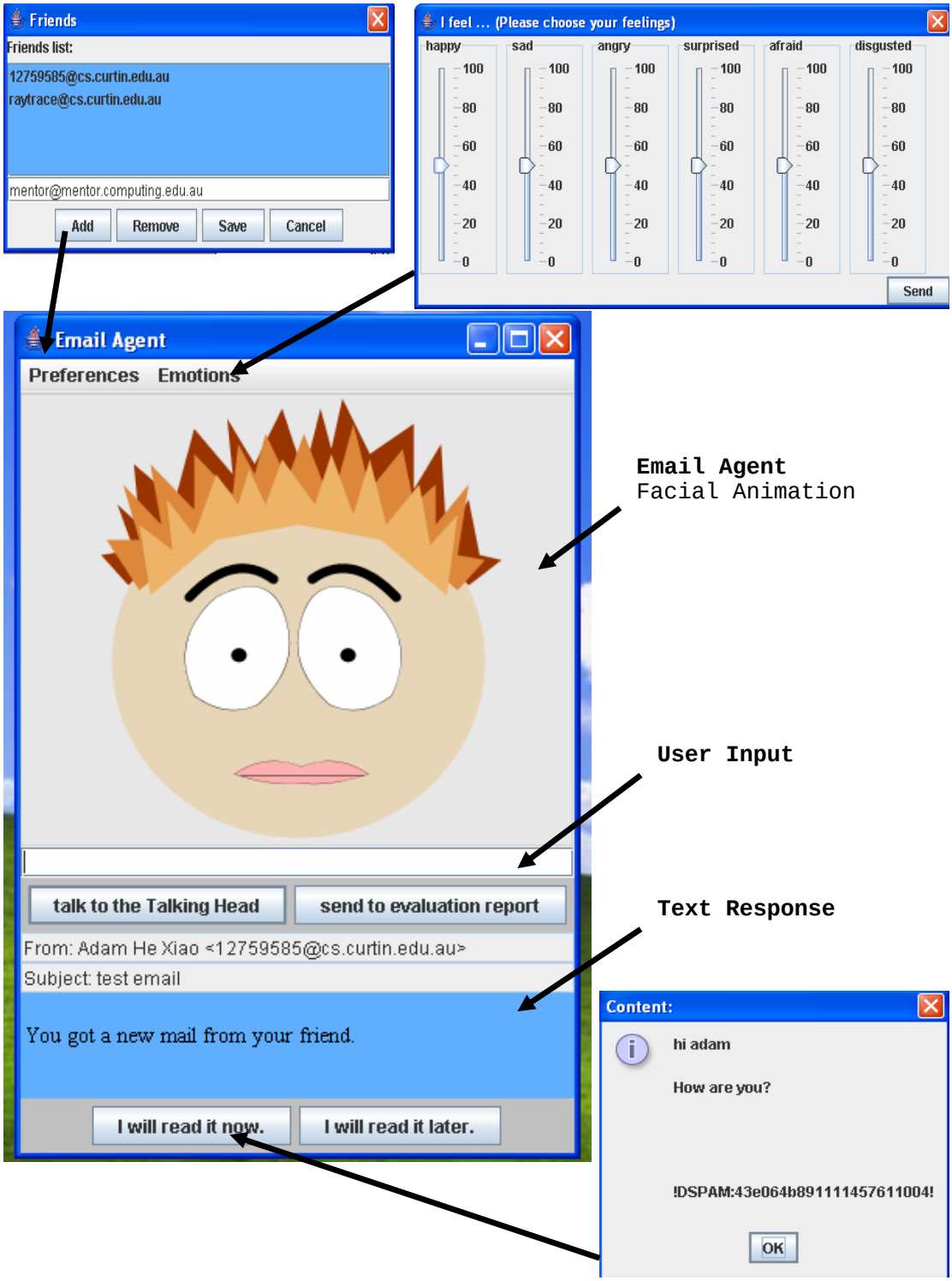


Figure 5.3: Screen Shot of Email Agent GUI

- The **"Friends" dialog** will pop up when the user presses the menu item "Friends" in the "Preferences" menu. The user can "add" new contacts or "remove" an email address from the friends list. The friends list will be saved as a file on the user's machine when the user presses the "save"

button. This information allows the personality to respond differently to friends.

- The **"Emotions" dialog** will pop up when the user presses the menu item "I feel" in the "Emotions" menu. The emotions, which reflect the user's feelings, will be saved to a server-side log file for later evaluation of the system by the researcher (see Figure 5.6).
- The **Email Agent drawing area** is an area where the ECA and its facial animation are shown. The size of the ECA will change when the user resizes the window. Note that for engagement, the ECA always looks at the user during interactions so the user can better perceive its personality. This feature depends on the ECA's personality. As Figure 5.4 shows, the ECA looks slightly left when it stays on the right of the screen and vice versa.
- The **User input field** is a text field to receive the user's conversational input.
- The **"talk to the Talking Head" button** indicates that the user wants to send their input as a question to the Mentor system.
- The **"send to evaluation report" button** indicates that the user wants to send their input to the evaluation log indicating information such as their feelings about the ECA's personality. The input will be saved to a server-side log file (see Figure 5.6) for evaluation.
- The **"From" field shows** where the mail comes from.
- The **"Subject" field** shows the subject of the mail.
- The **Text response field** is a text field showing what the ECA is saying. This is useful especially for evaluators with English as second language since they may not fully understand what the ECA says. And as for many public domain TTS synthesisers, the speech is not very high quality. The TTS uses Festival as the NLP, and Mbrola as the DSP, with the emotional content being inserted between these two stages (Stallo 2000).
- The **"I will read it now" button** can be pressed to pop up the email **"content" dialog** (see Figure 5.1).
- The **"I will read it later" button** can be pressed to clear the email content, "From" field and "Subject" field.

- The **Questionnaire dialog** (see Figure 5.5) will pop up after a user interacts with the ECA for several days. These questions belong to the formative evaluation, and relate to whether or not an ECA can choose its appearance and behaviour, and express its feelings according to its personality (see Chapter 6 for detail).

Figure 5.4: Screenshot of the Email Agent in Different Location in Screen

Figure 5.5: Formative Questionnaire

Additionally, usage data, evaluation reports and formative questionnaires are saved to the log file in the Mentor server (see Figure 5.6). The user's nickname is recorded instead of the user's real name since the evaluation is anonymous. The personality traits are also recorded since different evaluators may interact with ECAs with different personalities. The value "100" means that this ECA has personality A while "0" means it has personality B. A and B are two "opposite" personalities which were used to evaluate the personality model (see Section 6.2).

```
<logData date="Wed Feb 01 14:25:19 GMT+08:00 2006"
seconds="1138775119512">
nickname = adam    personalityTraits = 100
Type: user question to mentor
Content: hi
</logData>

<logData date="Wed Feb 01 14:25:19 GMT+08:00 2006"
seconds="1138775119542">
nickname = adam    personalityTraits = 100
Type: response from Mentor to user
Content: <p><happy>Good afternoon to you He.</happy></p>
</logData>
```

Figure 5.6: Example Data in Log File

5.5 Email Monitor

The Email Monitor module was developed using the JavaMail API, which is an object-oriented API to Internet messaging systems. The JavaMail API is designed to make it easy to add an electronic mail capability to simple applications (Sun Microsystems 1998). It includes appropriate convenience classes which encapsulate protocols such as Internet Message Access Protocol (IMAP), Simple Mail Transfer Protocol (SMTP) and Post Office Protocol (POP) (Sun Microsystems 1998).

The Email Monitor module has two parts:

- The first part makes a connection to the incoming mail server. The incoming mail server of Curtin's computing school is mail.computing.edu.au and it uses IMAP as protocol for receiving message. Classes from the JavaMail API for connecting to an IMAP server are used and the code is from examples in the JavaMail API Design Specification of Sun Microsystems, Inc (Sun Microsystems 1998). Currently the mail address is hard wired into the code but it can be altered to allow the program to run elsewhere.

- The second part keeps checking the mailbox to see if new mail has arrived. The mail monitor will check unread mail when the user first runs the Email Agent, and after that only new mail is checked. After the monitor receives a new mail it will send the sender's address and mail subject to the server as mentioned previously, and display the contents in the GUI.

After discussions within the research group, it was decided that users would not want to learn a new workflow for dealing with emails since all of them are highly experienced computing people and have their own habit of dealing with email. This was one of the main reasons for not developing a fully-fledged email client (See Section 7.4.1 for discussion).

5.6 Conclusion

This chapter discussed the third step of this research: implement the computational personality model within Curtin's ECA system. An experimental application - Email Agent - has been developed to implement the personality model and it runs within Curtin's ECA system. Section two discussed the system architecture, which includes the Mentor and VHML servers, server and client side personality model, etc. Section three, four and five presented the Email Agent program discussing its overall functionality, its user interface and its mail monitor module.

Chapter 6

Data Collection

6.1 Introduction

This chapter is concerned with the data collection of this research. The three steps of the experiment are introduced, followed by a description of the four parts of the questionnaire. The reasons for the questions are discussed in Section 6.3. Section 6.4 discusses the five point Likert Scale, as well as describing the statistical tests used to analyse the data given by such a scale.

6.2 Experiment

The questionnaire evaluation of the personality model was done by users who had interacted with the Email Agent over a long period of time. The purpose of the questionnaire was to get valuable feedback from users, and the feedback was then used to evaluate the hypotheses of this research. The implementation of the Email Agent has been discussed in Section 5.3.

Two ECAs were selected as email agents and they had "opposite" personality characteristics in terms of the FEN traits in Figure 4.1. Email Agent A had an extrovert, emotionally stable and friendly personality, while Email Agent B had an introvert, neurotic and hostile personality. The users who interacted with A were classified as belonging to group A while other users who interacted with B were classified as belonging to group B. Group A and B were supposed to have an equal number of users and their feedback was analysed separately.

The experiment was divided into three parts:

- Personal user information such as age, gender etc and user personality characteristics were recorded before the experiment. This information was obtained by the user filling in a demographic questionnaire that took roughly thirty minutes to complete. The standard 50 psychological questions from IPIP (Goldberg 2005) were used in this part to survey the users' personality characteristics.
- Then participants used the Email Agent for roughly two weeks. The program was invoked when a user logged in and should not have been terminated except by logging out. During this period, formative questions were popped up by the program several times. These questions were concerned with whether or not the ECA can choose its appearance, its behaviour and express its emotion according to its personality. The user's input and user's emotions were recorded in a log file and only identified by their nickname since the experiment was anonymous. The reasons why these questions were needed are detailed in Section 6.3.2.
- After the experiment, the user needed to complete the summative questionnaire, which related to the testing of hypotheses. There are two parts to the summative questionnaire; the first part concerns the user identification and description of the ECA's personality, and the second part concerns higher-level questions (see Section 6.6).

6.3 Questionnaires

The questionnaires were designed on the basis of the requirements of testing the personality model against the hypotheses. Table 6.1 shows the purpose of different questions and the related hypotheses that the questions wanted to test. Each question is described in its own subsection.

The question "Your nickname" is asked in each part of the questionnaire to identify the user. The nickname is recorded instead of the user's real name to make sure the participants are anonymous.

Table 6.1: Reason of Questions

Subsection	Purpose	Related Hypothesis
6.3.1 Demographic Questions (Online)	Questions concerning personal user information such as age, gender etc and user personality characteristics that can be used for classifying the user when analysing feedback.	
6.3.2 Formative Questions (Popped up by the program)	Questions concerning whether or not ECA can change appearance, behaviour and express his/her emotion according to his/her personality.	**Sub-hypothesis 1:** An ECA equipped with the personality model can express emotion according to its personality characteristics. **Sub-hypothesis 2:** A user can identify the personality or the personality characteristics of the ECA.
6.3.3 Summative Questions Part A (Online)	Questions concerning user identification and description of the ECA's personality.	**Sub-hypothesis 1:** as above **Sub-hypothesis 2:** as above
6.3.3 Summative Questions Part B (Online)	Higher level questions concerning whether this model enhances the believability of ECA and does it make human computer interaction more satisfying.	**Hypotheses:** The personality model will enhance the believability of ECA and make human computer interaction more satisfying.

6.3.1 Demographic Questions

The initial questions obtained the participants' demographic information. This demographic information can be used to group participants. Different groups may have different attitudes to the ECA with its personality. However, the more groups that are created, the smaller each group will be, assuming the number of participants is fixed. For example, "primary cultural or racial identification" has fourteen available groups and it needs a large number of participants. The distinction between feedback from each group will be small if the difference of each group is small. As most participants are computer science postgraduate students, their computing experience will be similar and it is useless to group them according to their computing experience.

The reasons for the questions in the demographic questionnaire (see Appendix B) are justified by the following:

- **Gender** is needed to classify the user since male and female may have different feelings to an ECA with a specific personality.
- **Age** is needed to classify the user since old and young users may have different feelings about the appearance of ECAs. For example, a young user may like "cool" ECAs.
- **First language** is necessary to classify the user since the ECA will speak English and a user with English as a second language may not understand it fully.
- **Primary cultural or racial identification** classifies the user's culture since different cultures may respond differently to ECAs. The cultural catalogue is from the standard psychological personality test found at http://www.outofservice.com/bigfive/
- **"How long have you used email?"** is used to group the participants since participants with different level of using email have different computing experience and may give different feedback. Long-term mail users will understand the application and may be likely to judge the personality while inexperienced users may confuse the email system difficulties with the personality.
- **"Have you ever used ICQ, MSN or similar system before?"** is used to classify the user since a user who likes to use ICQ or MSN may like this system as well.
- **"Have you ever used a Talking Head system before?" and "Have you ever used a Dialogue Management system before?"** is needed to classify the user since a user who has used a Talking Head system or Dialogue Management system before may be able to identify the ECA's personality faster and more accurately.
- **"Please use the rating scale below to describe how accurately each statement describes <u>you</u>."** is used to survey the participants' personality characteristics. This is needed to classify the participants for the analysis of their feedback. For example, a joyful participant may like a joyful ECA since their personality characteristics match. The 50 questions are from the standard

psychological personality test IPIP (Goldberg 2005). Figure 6.1 shows three sample questions from the fifty questions.

Figure 6.1: Sample Demographic Questions (see Appendix B)

6.3.2 Formative Questions

Formative questions were used to test, over time, the user's perception of the personality characteristics being modelled as well as test the sub-hypothesis 1: **An ECA equipped with the personality model can express emotion according to its personality characteristics**. The formative questionnaire (see Figure 5.5) was popped up by the Email Agent program several times during the experiment, asking the participant to use the Likert scale to describe how accurately each statement describes the ECA. The reasons for the formative questions are justified in Table 6.2.

Table 6.2: Reason for Formative Questions

Questions	Purpose
Does not talk a lot.	Tests whether or not the user can perceive an ECA's introversion or extroversion personality characteristics.
Likes to talk to a lot of different people.	
Likes changes.	
Gets upset easily.	Tests whether or not the user can perceive an ECA's Neuroticism and Emotion Stability characteristics.
Seldom feels blue.	
Is not really interested in others.	Tests whether or not the user can perceive an ECA's friendliness and extroversion characteristics.
Warms up quickly to others.	

6.3.3 Summative Questions

Table 6.1 on Page 67 shows the overall purposes of the two parts of the summative questionnaire. The purpose and reason for each question will be discussed in Table 6.3.

Table 6.3: Reason for Summative Questions (see Appendix D for questions)

Question Number	Purpose
Part A 1	Tests whether or not the user withdraws from the evaluation early and records their reason.
2 3 4 5	Question 2, 3, 4 and 5 are used to test the user's perception of the ECAs' characteristics of appearance, emotions, thoughts and behaviours.
6 7	Question 6 and 7 are needed to let us know how accurately the ECA's personality is identified or described by user.
Part B 8	Question 8 is used to obtain the user's attitudes to whether the ECA should have a personality.
9	This question is needed to let us know whether this model makes the interaction more satisfying. The user is asked "What kind of Talking Head would you prefer" which is useful for the future research of modelling suitable ECA's personality characteristics.
10	This question is needed to let us know whether this model enhances the believability of the ECA.
11	This question is needed to let us know whether the user felt engaged with the ECA.
12	Question 12 is used to test whether or not the scenario is suitable to evaluate the ECA's personality. The user also is asked to indicate other application that can be useful for the future research of design the long-term scenario.
13	Question 13 records any comments that the user wants to make.

6.4 Likert Scale

The Likert scaling approach (Burns & Bush 2000) is widely used for measuring the attitude of participants, in our case to the ECA. The Likert scale consists of a statement and numerically ranked categories of which a participant must choose the one option that best describes their perception of the statement (see Figure 6.1). An important part of the questionnaire for this evaluation is the standard five-point Likert scale that was used to measure perceptions of participants.

The statistical test performed is the t-test (Spiegel 1961). The normality assumption of the original data is obviously false. However, the purpose is to ensure the distribution of the mean is normal. As the data is in a fixed range, the original distribution has a finite variance, and so it is likely that even for 10 participants the inaccuracy due to the normality assumption is minor. Over 30 participants would be better.

With the Likert scale, the value 3 is considered to be neutral. The t-test was performed on the difference between the observed mean with the neutral value, to observe if the difference has any statistical significance within the 95% confidence interval. Thus, for each statistical test, two hypotheses were formulated as below (see Figure 6.2):

H_0 = there is **no difference** between the observed mean and the neutral mean
H_1 = there is a **significant difference** between the observed mean and the neutral mean

Figure 6.2: Statistical hypotheses

If the value 0 lies inside the 95% confidence interval of the difference, this will indicate that there is no difference or the difference is not significant at 95% confidence, thus resulting in the acceptance of H_0 and rejection of H_1. On the contrary, if the value 0 does not lie inside the 95% confidence interval of the difference, this will indicate that there is a significant difference at 95% confidence, thus resulting in the acceptance of H_1 and rejection of H_0. The 95% confidence interval of the difference in SPSS indicates the difference between observed means and neutral value 3. So the value 0 actually means 3 + 0 = 3.

Following the analysis, if the difference was significant and positive, then it could be concluded that the observed mean was significantly better than the neutral mean. On the contrary, if the difference was significant and negative, then it could be concluded that the observed mean was significantly worse than the neutral mean. The analysis was performed using SPSS and will be discussed in Chapter 7.

It is important to note that the evaluation results will not be highly accurate if the sample size is very small. The small sample size is a considerable problem of this evaluation of an ECA with personality (see Section 7.2).

6.5 Conclusion

This chapter discussed the data collection of this research.

- Section 6.2 introduced the experiment being used for testing the hypothesis.
- Section 6.3 discussed the reasons for questions in the questionnaire.
- Section 6.4 discussed the five point Likert Scale and the t-test used to analyse the attitude.

Chapter 7

Analysis and Results

7.1 Introduction

The previous chapter described the experiment and the data collection of this research. This chapter details the data analysis and results of the evaluation. Participant demographic information is discussed at first, as well as the problem of the small sample size. An analysis of the participants' feelings and formative questions is discussed in Section 7.3. Then the summative questions are analysed one by one through analysis of the Likert scales data and qualitative data. Finally, the problems that affected the evaluation are discussed in Section 7.5.

7.2 Participant Demographics

There were 10 participants involved in the evaluation. All of them were postgraduate students (Honours, Masters and PhD students) or staff members of the Department of Computing. A sample size of 10 was not satisfactory for evaluating the ECA experiment. However, several difficulties were encounters when looking for volunteers.

- The Email Agent connected to servers in the Department of Computing (VHML and Mentor servers), as well as to the mail server. This required that participants had valid computer accounts in the Department of Computing and that meant that the participant should be students or staff in the department.
- The evaluation period was during the university's summer holiday. There were no undergraduate students in the department and some postgraduate students and staff were on holiday. Since only postgraduate or staff members were available, the sample size was quite small.

- Some people refused to join the evaluation since it was a long-term evaluation and needed around two weeks of daily interaction.
- The application included animation and audio, and some potential volunteers withdrew as they mistakenly believed it would consume too much of their system memory and affect the speed of their machine.
- Some persons refused to join the evaluation since the evaluation needed to measure their personality and they would have had to fill in the 50 psychological questions.

```
I don't want to use the agent as I don't want it intruding into my
work, requiring me to look after it when I am doing other things.

I don't like the distracting movement or the fact that the window
consumes screen space.
```

Quote 7.1: Comments of a User Who Refused to Evaluate the System

The above quote came from a user who refused to evaluate the Email Agent. This indicated the difficulties for recruiting participants. The sample size is a considerable problem of this evaluation of an ECA with personality. Obviously quantitative results will not be highly accurate. In order to improve the accuracy of the results, analysis of qualitative feedback was considered to be necessary. All participants were encouraged to give as much qualitative feedback as possible and the analysis of their comments will be described in Section 7.4.

Figure 7.1 (a), (b) and (c) shows the bar charts of the demographic statistics.

- As Figure 7.1(a) shows, there were 8 male and 2 female participants in the evaluation, which is a standard distribution in the Department of Computing.
- Figure 7.1(b) indicates that 6 participants had never used a Talking Head or Dialogue Manager before.
- The "years of use email" shown in Figure 7.1(c) reflects the user experience of using computers. Most of the participants had used email for more than 8 years, indicating a high level of computing experience. This is because they are either computer science postgraduate students or computer science staff. This indicated another problem: computing persons may not share the views of non-computing person. This evaluation only involves high level computing people (see Quote 7.2). An ideal evaluation should involve both computing and non-computing person and their attitudes need to be compared.

```
Probably since I am a computing person and so don't need to have
the interface humanised for me. However, I am aware that there is
a (large) target audience that doesn't share my views.
```

Quote 7.2: Comment of the Question "Why do you feel that a Talking Head needs or does not need a personality?"

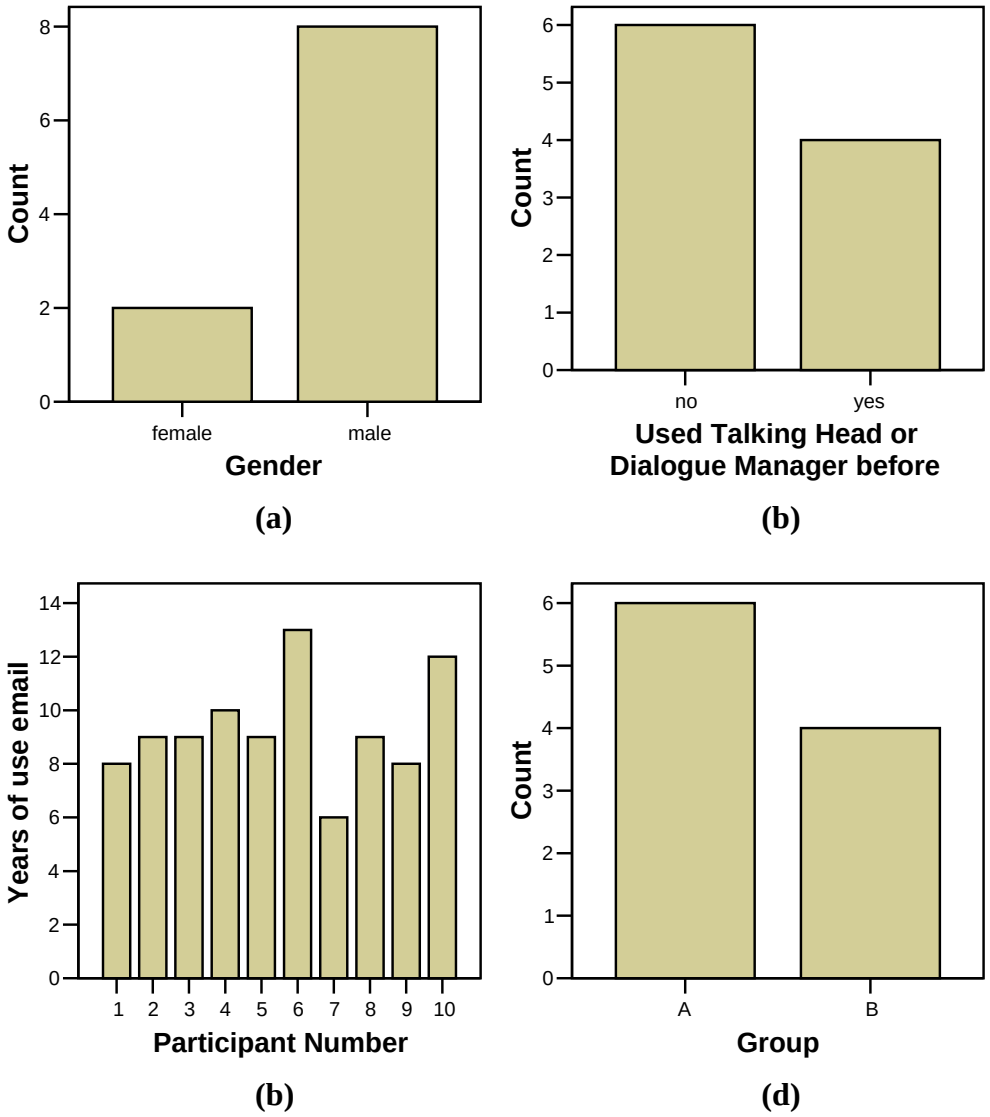


Figure 7.1: Demographic Information (a) Gender (b) Used Talking Head or Dialogue Manager before (c) Years of Use Email and (d) Group

Figure 7.1(d) shows that 6 participants interacted with Email Agent A while 4 participants interacted with Email Agent B. Section 6.2 indicated that group A and B were supposed to have equal number of users. At first, the same number of Email Agent A and B were installed according to the requirement. However, some

participants dropped out either because they changed their mind or because they wanted the system to work at home.

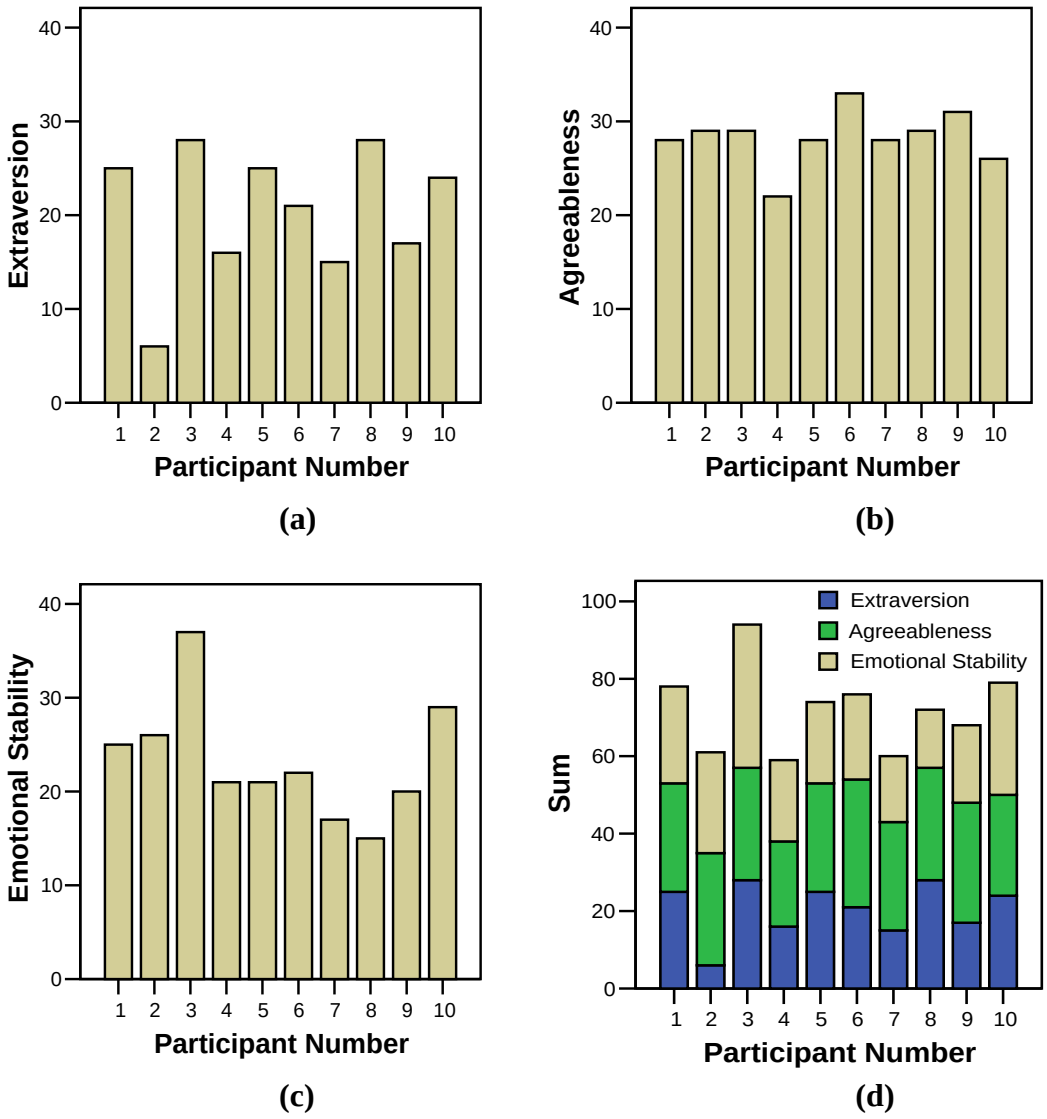


Figure 7.2: Participant Personality Traits (a) Extraversion (b) Agreeableness (c) Emotion Stability and (d) Sum of (a) (b) and (c)

Participants' personality traits were measured by the standard psychological questionnaire and were used to classify the participants for the analysis of their feedback. The values for the Big Five traits of Extroversion, Agreeableness, Conscientiousness, Imagination and Emotional Stability, were calculated from the answers of the 50 psychological questions (Goldberg 2005). The traits presented in Figure 7.2 only cover *Extroversion, Agreeableness and Emotion Stability* – these being the traits relevant to the study.

In Figure 7.2, the value “40” is an extremely high score and “0” is an extremely low score. The value “20” can be considered as a neutral value. It can be seen from Figure 7.2(a) that participant number 2 rates “6”, which implies this participant is a very introverted personality. Others are around “20”, and that means that they have either a slightly extroverted or slightly introverted personality. It will be seen in Section 7.4 that the feedback from participant number 2 compared with the feedback from the other participants is different. Figure 7.2(d) shows the piled values of these three traits. Participant number 3 got the highest pile and his specific feedback is also compared to others. This data “classified” each participant’s personality.

7.3 Formative Evaluation

7.3.1 Emotions

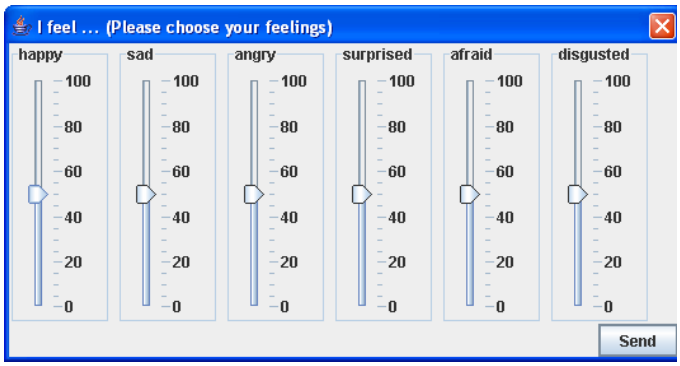


Figure 7.3: Emotions Dialog

The emotions dialog (see Figure 7.3) was used to record the user’s feelings during the interaction and the values were sent to the server-side log file when the user pressed the “send” button. A total of 12 emotion dialog interactions were logged during the two-week evaluation. There were 4 feedbacks from group A and 8 from group B.

The users’ feelings about the interaction were expected to be:

- Expectation 1: Group A should happier than group B since Email Agent A had an extrovert friendly personality.
- Expectation 2: Group B should feel sad, angry or disgusted since Email Agent B had a hostile personality.

- Expectation 3: Group A and B should have “surprised” or “afraid” values around 50. A value of 50 is neutral, and feelings of surprised and afraid should be neutral since the ECA’s personality characteristics do not indicate these feelings.

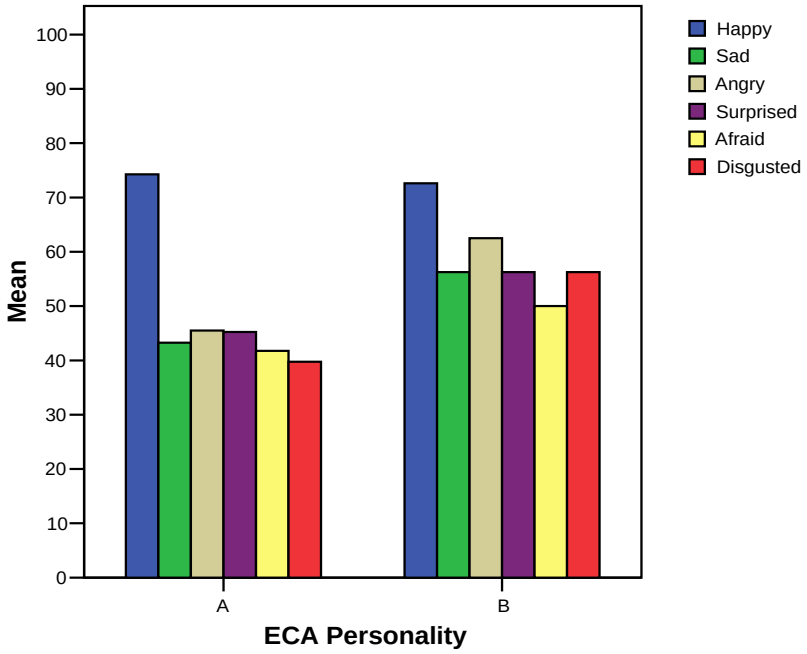


Figure 7.4: Means of Emotion Intensities of Participants in Group A and B

Table 7.1: Statistics of Means of Emotional Intensities

	Group	N	Mean	Std. Deviation	Std. Error Mean
Happy	A	4	**74.25**	16.276	8.138
	B	8	**72.63**	24.974	8.830
Sad	A	4	**43.25**	13.500	6.750
	B	8	**56.25**	17.678	6.250
Angry	A	4	**45.50**	9.000	4.500
	B	8	**62.50**	23.146	8.183
Surprised	A	4	**45.25**	9.500	4.750
	B	8	**56.25**	17.678	6.250
Afraid	A	4	**41.75**	16.500	8.250
	B	8	**50.00**	**.000**	**.000**
Disgusted	A	4	**39.75**	20.500	10.250
	B	8	**56.25**	17.678	6.250

Figure 7.4 shows the Means of emotional intensities of participants in groups A and B. The values of Means can be seen in the Table 7.1. In Table 7.1, “N” indicates the number of feedback responses instead of the number of participants. Some participants sent more than one emotion intensity feedback, whilst some participants

sent none. Before presenting the results, it is necessary to stress that the number of responses - 4 and 8 - may be not solid enough to support or reject the expectation.

- Expectation 1 is not supported strongly by the result. Participants in Group A felt happy (mean = 74.25) when they interacted with the friendly ECA. However, participants in Group B also felt happy and the value is similar to Group A (72.63 compared to 74.25). The reason may be that some participants in Group B still felt happy when the hostile agent notified them when new mail arrived. Even a hostile agent has some satisfaction function.
- Expectation 2 is supported by the results since participants in Group B have a higher Mean than Group A: Sad 56.25 compared to 43.25; Angry 62.5 compared to 45.5 and Disgusted 56.25 compared to 39.75.
- Expectation 3 is supported by the data. Group A and B have "surprised" or "Afraid" values not far away from 50 and Group B have higher values than Group A. The Mean of the "Afraid" value of Group B is 50 and the Standard Deviation is 0. That indicates that none of them had a feeling of being afraid.

Due to the small sample size, a formal break up by personality of participants as well as by ECA personality will not be informative.

7.3.2 Formative Questionnaire

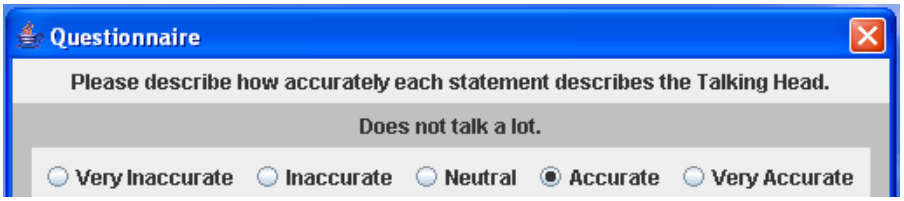


Figure 7.5: Formative Questionnaire Dialog (see Figure 5.5)

The formative questionnaire dialog (see Figure 7.5) included questions that relate to whether or not the ECA can choose its behaviour and express its feelings according to its personality. 15 feedback responses were received during the two-week evaluation. There were 7 feedbacks from group A and 8 from group B. Some participants answered the questions more than once and hence 15 feedbacks instead of 10 (the number of participants).

The user's answers of the questions were expected to be:

- Expectation 1: Group B should agree with "Does not talk a lot" since Email Agent B is an introvert, while Group A should disagree since Email Agent A is an extrovert. Group A should agree with "Likes to talk to a lot of different people" while Group B should disagree.
- Expectation 2: Group B should agree with "Get upset easily" since Email Agent B is neurotic and hostile while Group A should disagree since Email Agent A is emotionally stable. Group A should agree with "Seldom feels blue" while Group B should disagree.
- Expectation 3: Group B should agree with "Is not interested in others" since Email Agent B is introvert and hostile while Group A should agree with "Warms up quickly to others" since Email Agent A is extrovert and friendly.
- Expectation 4: Group A should agree with "Likes changes" since Email Agent A has the personality characteristic of liking to change his appearance.

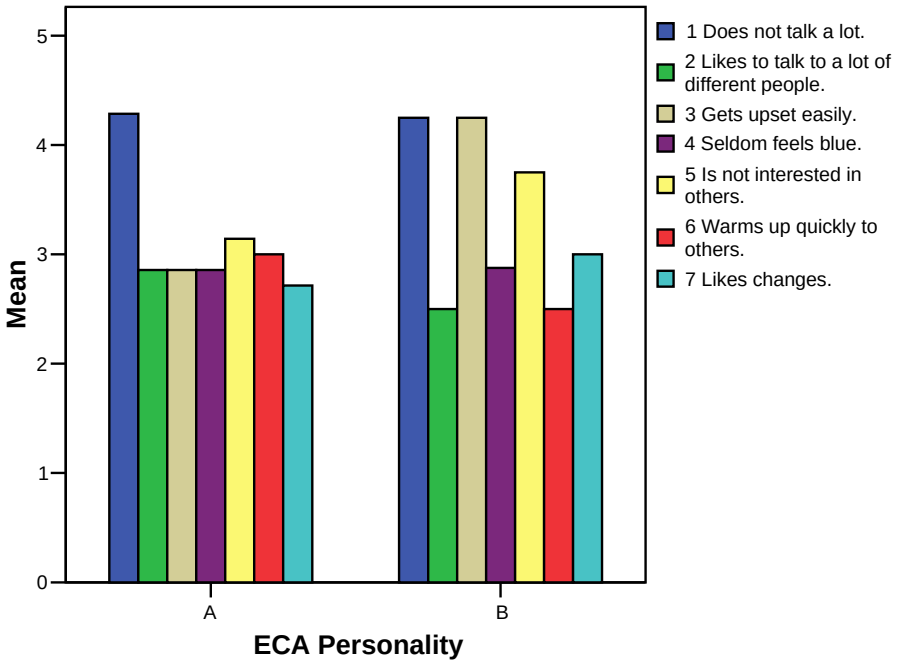


Figure 7.6: Means of the Likert Scales of the Formative Statements of Participants in Group A and B

Figure 7.6 shows the Means of each statement described by participants. According to the expectations, the Means of statements 1, 3 and 5 by Group A are supposed to be significantly lower than the means by Group B, while other statements are supposed

to be significantly higher. As Figure 7.6 shows, only statements 3 and 5 by Group A are obviously lower than B. Both Groups agree with the statement 1. The data was further analysed using statistical tests, and the possible reasons for this issue are discussed.

7.3.2.1 Expectation 1

Figures 7.6 and Table 7.2 show that statement "Does not talk a lot" is agreed to by both Group A and Group B since the Means are 4.29 and 4.25. And the answers are supported by the t-Test as significantly different from the neutral value 3 since the *95% Confidence Interval of the Difference* is 0.41 to 2.17 and 0.38 to 2.12 respectively and these do not cover 0 (see Table 7.3). However, Group A is supposed to disagree with the statement by the expectation. The reason may be the frequency of active responses is seen to be very slow (one time each eight hours) even for the extrovert ECA, and that may make the user think: "It didn't talk much".

Table 7.2: One-Sample Statistics of the Means of the Likert Scales of statements "Does not talk a lot" and "Likes to talk to a lot of different people"

	N	Mean	Std. Deviation	Std. Error Mean
Does not talk a lot (Group A)	7	**4.29**	.951	.360
Does not talk a lot (Group B)	8	**4.25**	1.035	.366
Likes to talk to a lot of different people (Group A)	7	**2.86**	.900	.340
Likes to talk to a lot of different people (Group B)	8	**2.50**	.926	.327

Table 7.3: One-Sample t-Test of the Means of the Likert Scales of statements "Does not talk a lot" and "Likes to talk to a lot of different people"

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Does not talk a lot (Group A)	3.576	6	.012	1.286	**.41**	**2.17**
Does not talk a lot (Group B)	3.416	7	.011	1.250	**.38**	**2.12**
Likes to talk to a lot of different people (Group A)	-.420	6	.689	-.143	-.97	.69
Likes to talk to a lot of different people (Group B)	-1.528	7	.170	-.500	-1.27	.27

The frequency of active responses is defined to be very low to prevent it annoying participants and affecting their work, otherwise more participants may drop out before the two weeks duration of the experiment. This problem also affected the accuracy of

the users' identification of extrovert and introvert in the summative questions. The answers for the statement "Likes to talk to a lot of different people" are seen to have no significant difference with the neutral value (see Table 7.3). Therefore, Expectation 1 is partially supported by the evaluation results.

7.3.2.2 Expectation 2

Tables 7.4 and 7.5 show that the statement "Get upset easily" is agreed to by Group B since the results are proved by the t-Test as significantly higher than the neutral value 3. The lower and upper values of 95% confidence interval of the difference are 0.86 and 1.64 and this does not cover 0. The Mean is also higher than the Mean of Group B (4.25 compared to 2.86). This result supported expectation 2 but it must again be pointed out that the sample size is not solid enough.

Table 7.4: One-Sample Statistics of the Means of the Likert Scales of statements "Get upset easily" and "Seldom feels blue"

	N	Mean	Std. Deviation	Std. Error Mean
Get upset easily (Group A)	7	2.86	1.069	.404
Get upset easily (Group B)	8	**4.25**	.463	.164
Seldom feels blue (Group A)	7	2.86	1.069	.404
Seldom feels blue (Group B)	8	2.88	1.126	.398

Table 7.5: One-Sample t-Test of the Means of the Likert Scales of statements "Get upset easily" and "Seldom feels blue"

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Get upset easily (Group A)	-.354	6	.736	-.143	-1.13	.85
Get upset easily (Group B)	7.638	7	.000	1.250	**.86**	**1.64**
Seldom feels blue (Group A)	-.354	6	.736	-.143	-1.13	.85
Seldom feels blue (Group B)	-.314	7	.763	-.125	-1.07	.82

As Table 7.5 shows, the answers for the statement "Seldom feels blue" by both Group A and B are proved to be having no significant difference with the neutral value. Informal feedback from a participant indicated that the reason was that the participant did not understand the question fully. The same problem may occur in other parts of the questionnaire, since 3 of the 10 participants did not have English as their first

language. Unfortunately, giving more explanation of terms and questions runs the risk of biasing the response from the user, or pre-conditioning them to a preferred answer. Expectation 2 is partially supported by the evaluation results.

7.3.2.3 Expectation 3

Tables 7.6 and 7.7 show the statement "Is not interested in others" is agreed to by Group B, since the results are proved by the t-Test as significantly higher than the neutral value 3. This result supports expectation 3 since an introvert and hostile personality should be recognized as "not interested in others". However, the answers to the statement "Warms up quickly to others" by both Group A and B are proved to be having no significant difference with the neutral value. The reason may be that the feedback from this statement is affected by the incorrect recognized statement "Does not talk a lot". As they determined it did not talk a lot, they did not discover what it said of others and so indicated neutral to "Warms up quickly to others".
Expectation 3 is partially supported by the evaluation results.

Table 7.6: One-Sample Statistics of the Means of the Likert Scales of statements "Is not interested in others" and "Warms up quickly to others"

	N	Mean	Std. Deviation	Std. Error Mean
Is not interested in others (Group A)	7	3.14	.900	.340
Is not interested in others (Group B)	8	**3.75**	.707	.250
Warms up quickly to others (Group A)	7	3.00	.577	.218
Warms up quickly to others (Group B)	8	2.50	.756	.267

Table 7.7: One-Sample t-Test of the Means of the Likert Scales of statements "Is not interested in others" and "Warms up quickly to others"

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Is not interested in others (Group A)	.420	6	.689	.143	-.69	.97
Is not interested in others (Group B)	3.000	7	.020	.750	**.16**	**1.34**
Warms up quickly to others (Group A)	.000	6	1.000	.000	-.53	.53
Warms up quickly to others (Group B)	-1.871	7	.104	-.500	-1.13	.13

7.3.2.3 Expectation 4

Expectation 4 was rejected by the analysis of data for the following reasons:

- The answers to the statement "Like changes" by both Group A and B are proved to be having no significant difference with the neutral value (see Table 7.8).
- The Mean of Group B is higher than Group A (3.00 compared to 2.71) while it is supposed to be lower than Group A.

The reason, found by a short informal interview, was that participants misunderstood the meaning of the statement. Some of them thought that the statement was asking about "Like change their emotions" instead of "Like change their appearance". The Email Agent A is designed to like to change his appearance while Email Agent B likes to change his emotions or moods. The statement "Likes changes" is a vague statement although it used by the standard psychological description.

Therefore, none of these expectations are fully supported by the evaluation results due to lack of experience of evaluating personality of an ECA. However, this pilot study is valuable in that it has indicated problems that may affect future studies. Again, the results are not very accurate due to the small sample size.

Table 7.8: One-Sample Statistics of the Means of the statement "Likes changes"

	N	Mean	Std. Deviation	Std. Error Mean
Like changes (Group A)	7	**2.71**	1.254	.474
Like changes (Group B)	8	**3.00**	.535	.189

Table 7.9: One-Sample t-Test of the Means of the statement "Likes changes"

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Like changes (Group A)	-.603	6	.569	-.286	**-1.45**	**.87**
Like changes (Group B)	.000	7	1.000	.000	**-.45**	**.45**

7.4 Summative Evaluation

7.4.1 How long did you use the Email Agent system?

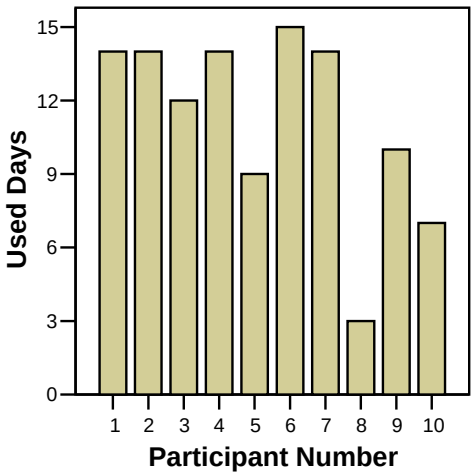


Figure 7.7: Number of Days Each Participant Used the Email Agent System

It can be seen from the Figure 7.7 that most participant used the system for about two weeks. Quote 7.3 contains two quotes, one by participant 5, and the other by participant 10, as to why these participants left the evaluation. They indicated two shortcomings of the system: Email cannot be replied to in the same interface, and the system does not support the viewing of email attachments. A feature for answering email was not considered as a necessary function because it was decided that users would not want to learn a new workflow for dealing with emails since all of them are highly experienced computing people and have their own habit of dealing with email.

The system functionality was similar to the old “you have mail” notification that used an icon on the desktop of a mailbox with the flag either up or down. In this TH version, the system will talk to the user about the title and sender of the mail. Also in future studies, since the system is component based, it can be integrated into an existing mail reader such as Outlook Express, Evolution or Thunderbird.

Participant number 8 only used the system for 3 days before a change of office occurred. As the log of that participant’s interaction indicates, a large number of interactions had taken place; the data gathered from that participant has been retained.

5: I got sick of using it after this time, mostly because I could not reply to email messages in the same interface. This would be a nice feature.
10: I frequently received emails with attachments, so I didn't interact with the program a lot.

Quote 7.3: Comments of a User Who Refused to Evaluate the System

7.4.2 User Perception of ECAs' Personality Characteristics

7.4.2.1 Appearances

Email Agent A and B had the same basic physical appearance. The same basic appearance should be set for ECAs with different personalities otherwise the radically different appearances will affect the user's perception of the ECAs personality characteristics and that may make the experiment inaccurate. However, minor changes to the appearance may occur because of the ECA's personality. For example, an extrovert ECA may want to change his/her hair colour more often.

- Expectation of appearances characteristics of Email Agent A and B: Email Agent A does not care about its appearance and likes to change its hair and lip colour. Email Agent B does not like to change its appearance.

As Quote 7.4 shows, the characteristics of "change its hair colour" and "does not care about his appearance" is recognized by participant 7 and 9. However, Email Agent A was described by participant number 8 as "looking quite serious", which is not expected by the personality model. Similar descriptions can also be found in Quote 7.6: Email Agent A looks unhappy and sometimes frowns. The reason may be that the physical appearance looks a bit unhappy and serious (see Figure 7.8).

3(B) Fairly bland, but at least it's not annoying (like the MS Office clip) and the idle-time animations are enough to ensure it isn't 'boring'.
7(A) it **don't care about his appearance**, I think.
8(A) It looks **quite serious** most of the time.
9(A) It **often changes its hair colour.**
10(A) Seems pretty good for what it is. A 3D version would be nice too.

Quote 7.4: Comments of the Question "Please describe the characteristics showed by the Talking Head concerning his/her <u>appearances</u> (for example you may answer "It always makes itself look cool")**."**

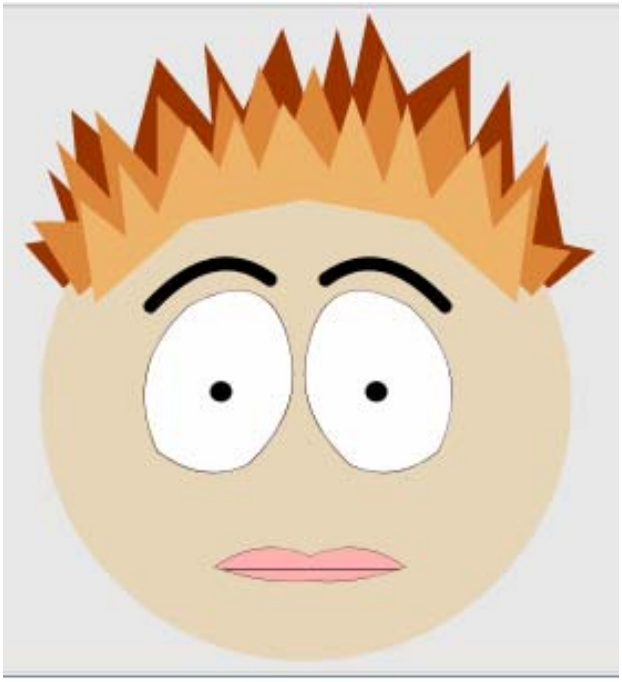


Figure 7.8: Physical Appearance of Email Agent A and B

1(B) It's **disgusting.**
3(B) **Rarely talks** - only for mail messages as far as I could tell. This is fine by me because I am fairly utilitarian when it comes to computers - I find unnecessary interruptions by a computer agent annoying.
4(B) **Doesn't talks a lot**, just blinks.
5(A) It **moves around lots.** This can be somewhat **distracting.**
6(B) **Doesn't move much nor talks a lot.**
8(A) It likes to **bob its head a lot.**
9(A) **It didn't talk much.** Actually, it did not begin any conversations. **Very polite** but seldom speaking.
10(A) I thought it **moved around too often** and was somewhat **distracting.** The constant **head rotation and eye movements** blinking were a bit too much.

Quote 7.5: Comments of the Question "Please describe the characteristics showed by the Talking Head concerning his/her <u>behaviours</u> (for example you may answer "It always talks a lot and keeps moving").”

7.4.2.2 Behaviours

- Expectation of behaviour characteristics of Email Agent A and B: Email Agent A likes to talk and move, and its behaviour is friendly. Email Agent B does not talk a lot and does not keep moving. Its behaviour is hostile and disgusting.

As Quote 7.5 shows, the characteristics of “likes to move” and “does not like to move” are recognized correctly by most participants. The characteristic of “does not

talk a lot" is confirmed by Group B as correct. The characteristic of "likes to talk" is denied by participant 9 who said: "It didn't talk much". The reason for this was discussed in Section 7.3.2.1 when analysing the feedbacks to the statement "Does not talk a lot". Participants 5 and 10 indicated that Email Agent A is somewhat distracting which is reasonable since A has an extreme extrovert personality. Participant 1 perceived Email Agent B as "disgusting" and participant 9 perceived Email Agent A as "very polite". This supported the expectation that Email Agent A has friendly behaviours and B has hostile behaviours.

Email Agent A's extremely extroverted personality behaviours were recognized correctly by participant 10. Participant 10 state that it "moved around too often", "The constant head rotation and eye movements blinking were a bit too much" which is exactly what is modelled by the personality of Email Agent A.

7.4.2.3 Emotions

- Expectation of emotional characteristics of Email Agent A and B: Email Agent A is always happy and does not like to change its emotions and moods. Email Agent B likes to have the emotions of angry and disgust, and likes to change its emotions and moods.

1(B) **disgust**
6(B) Generally **pretty angry/annoyed**?
7(A) It always **looks unhappy**.
8(A) It sometimes **frowns**.
9(A) I thought it was **happy**.

Quote 7.6: Comments of the Question "Please describe the characteristics showed by the Talking Head concerning his/her <u>emotions</u> (happy, angry, disgust, sad, surprise etc)."

As Quote 7.6 shows, the emotions of happy, disgust and angry are perceived the same as what is expected (see participant 1, 6, and 9 in Quote 7.6). The reason why Email Agent A is described by participant 7 and 8 as unhappy and frowns is discussed in Section 7.4.2.1. However, 4 of the 10 participants could not discern any emotions from the Email Agent. That may mean that the experimental scenario of email

notification is not good enough to show an ECA's emotions, and this problem will be discussed in the Section 7.5.2.

7.4.2.4 Thoughts

- Email Agent A is expected to be self-controlled and optimistic while B is expected to have low self-control and pessimism.

Quote 7.7 shows that Email Agent A is described as self-controlled by 8 and as optimistic by 9. Email Agent B is described by 1 and 6 as pessimistic. The reason why Email Agent A is described by participant 7 and 8 as pessimistic may be because participant 7 and 8 think that Email Agent A is unhappy and sometimes frowns (see Quote 7.6). Four of the ten participants could not discern any thoughts from the ECA. This is not surprising since thoughts have been indicated as very difficult to model and be perceived. The participants' answers concerning self-control, optimism or pessimism may have been lead or pre-conditioned by the questions since the questions already named these characteristics (see the caption of Quote 7.7).

1(B) **pessimism**
2(A) I couldn't discern any apparent thoughts.
3(B) Well, the few times I played around with the 'Talk to the head' option, it quickly devolved into unintelligible 'dialogue' where its responses didn't make much sense in the context of the dialogue. Not surprising considering how hard natural language processing is. So I can't say I was left with the illusion that it had 'thoughts' because it wasn't able to maintain the illusion of dialogue for long.
5(A) If the talking head had thoughts, I did not realise.
6(B) Along the lines of the above comment, I guess you would call it **pessimistic**.
7(A) **pessimism, a bit aggressive.**
8(A) It seems to have **a lot of self-control**, but tends to be **slightly pessimistic**, especially about the weather.
9(A) **optimism**

Quote 7.7: Comments of the Question "Please describe the characteristics showed by the Talking Head concerning his/her <u>thoughts</u> (self-control, optimism or pessimism etc)."

7.4.2.5 Personality

Quote 7.8 shows the description of the ECAs' personality by Group A and B. Blue text represents characteristics that are recognized correctly and the red text represents

characteristics that are perceived opposite to the personality model. The red text which all comes from Group A either because they believe Email Agent A "Does not talk a lot" or because they think Email Agent A has an unhappy appearance (see Section 7.3.2.1 and 7.4.2.1). Statements by participants 5 and 9 indicate the Email Agent A was "very quiet" and it is assumed that this was because of the low frequency of active responses that was set for both of the ECAs.

1(B) **not friendly, get bored and inpatient so easily** no matter what you are trying talking to him
5(A) The talking head on my machine was **very quiet**. It only informed me about email when it arrived and little else.
6(B) **Mostly quiet.** When not quiet, **engaged in outbursts about its dislike for me. Responded occasionally to a nice word,** with a thanks or something similar.
7(A) **aggressive, not friendly.**
8(A) I feel that the Talking Head has a lot of personality, it simulates the actual persona of a human quite well. I think the way it tells jokes is very good, really brightens my day sometimes. In my opinion it **looks a little too serious** and can do with more expressions. But overall I think it is an intriguing character.
9(A) It is **friendly** but **very quiet**.
10(A) **Nervous twitch**.

Quote 7.8: Comments of the Question "Please describe the personality of the Talking Head in your own words."

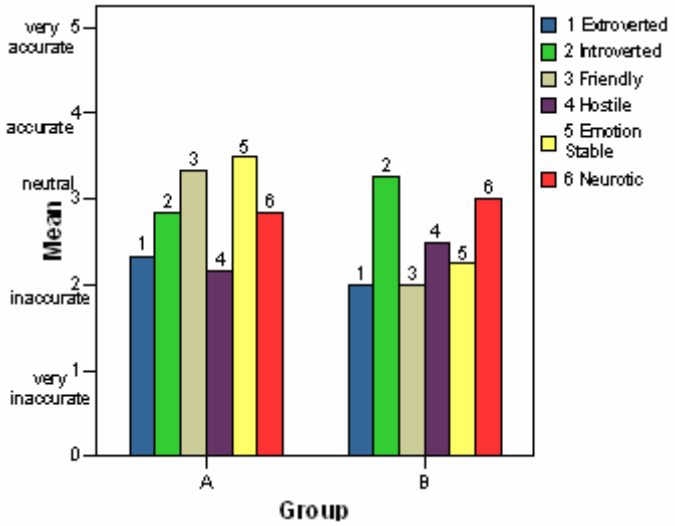


Figure 7.9: Means of Likert Scales of Personality Traits Answered by Participants for the Question "Please use the rating scale below to describe how accurately each personality label describes the Talking Head"

Email Agent A has the personality traits of extroverted, friendly and emotionally stable while Email Agent B has the personality traits of introverted, hostile and

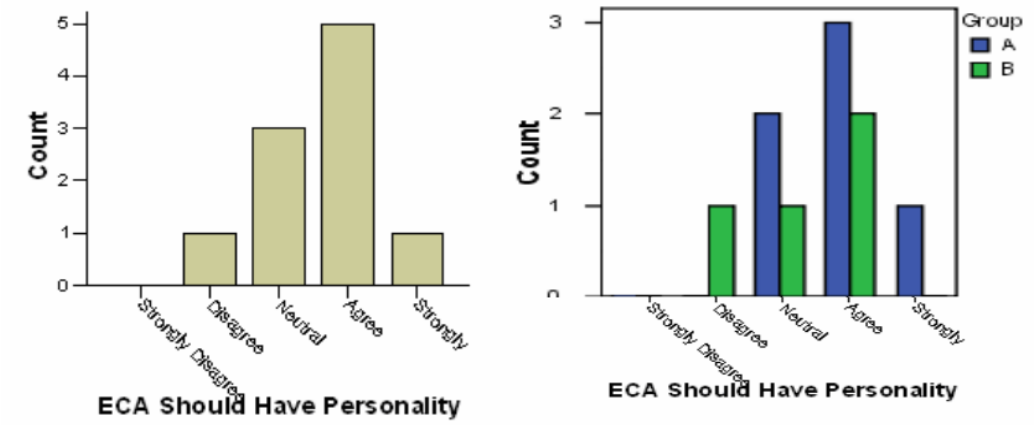
neurotic. As Figure 7.9 shows, the Means of the values for extroverted, friendly and emotionally stable of Group A are higher than that of Group B, while the Means of the other three traits are lower than Group B. This indicates participants can rank the personality traits correctly. However, users are not sure exactly which personality traits Email Agent has, since all the Means are around the neutral value.

7.4.3 High Level Questions Concerning User's Attitudes

7.4.3.1 Do you agree that a Talking Head should have personality?

2(A) Depends entirely on what the talking head is being used for.
3(B) Personally, I see no reason to try to add personalities to computers. Probably since I am a computing person and so don't need to have the interface humanised for me.
4(B) Depends on the personality of the talking head and the personality of the user.
5(A) They could be used to make an application easier to use - particularly for people who are not familiar with computers or certain programs.

Quote 7.9: Comments of the Question "Do you agree that a Talking Head should have personality?"



(a) (b)

Figure 7.10: Distribution of Results of the Likert Scale for the Question "Do you agree that a Talking Head should have personality?" (a) All Participants (b) Group A vs. Group B

Figure 7.10(a) shows that 6 of 10 participants agreed with this statement while only 1 participant disagreed. Figure 7.10(b) indicates that participants in Group A tend to agree with the statement and the participant who disagreed with the statement was in Group B. The ECA's personality may affect the answer to this question. However, the result may also be inaccurate due to the sample size being too small. Useful comments for this question are shown in Quote 7.9. They indicate that whether an

ECA should have personality depends on the scenario (Quote 7.9 2), the user's personality and the ECA's personality (Quote 7.9 4).

Although the quantitative results strongly support the proposition that an ECA should have a personality, the qualitative feedback in Quote 7.9 is not as strong. Quote7.9 (3) indicated that computing people may not share the views of non-computing people. In order to obtain a solid result to this question, future evaluation should involve large number of both computing and non-computing people. The comparison of answers of these two groups may be useful since high-level computing people tend to use professional applications, while non-computing people tend to use applications that they are not totally familiar with. The personality-rich ECAs may be required by non-computing people since these ECAs "could be used to make an application easier to use - particularly for people who are not familiar with computers or certain programs" (see Quote 7.9 5).

This issue, and some of the following results from the research, seem at odds with current thinking on ECAs. These issues are discussed in Section 7.5.

7.4.3.2 How was your satisfaction with the interaction with the Talking Head?

Figure 7.11 and Quote 7.10 shows users' attitudes to interaction with the ECA. Participant 2 said: "I only discovered one question it would actually answer" while participant 7 said: "at least it answers my questions". One reason for this difference may be found in the demographic information: participant 7 had experience of using the Mentor system while participant 2 had never used it before. Another reason may be that participant 2 did not want to try more questions since he has an extreme introvert personality. This feedback indicated some problems with the evaluation such as more domain knowledge being needed, and that participants needed to be trained about currently available topics before they can use it. Comments from participant 3, 6, 8 and 9 indicated the personality characteristics (blue text) that are expected by the personality model.

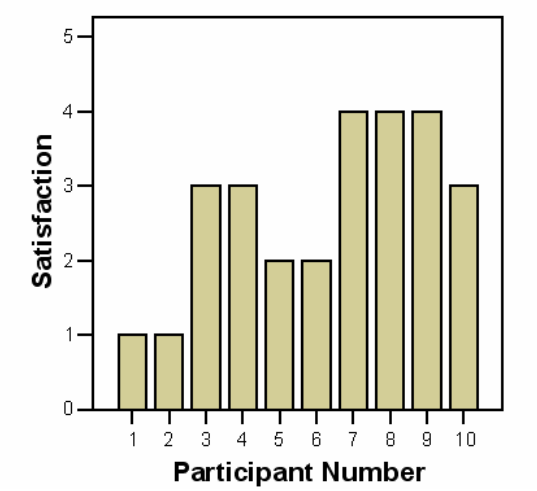


Figure 7.11: Results of the Likert Scale for the Question "How was your satisfaction with the interaction with the Talking Head?"

2(A) Almost everything I said seemed to confuse it. **I only discovered one question it would actually answer** ("what's the time?"), and the answer was not useful.
3(B) It did its job (**tell me when e-mail occurred) and didn't annoy** me beyond that. But then, a beep would do the same job too!
6(B) Couldn't make much headway with it. For example, why was **it angry most of the time?** I.e. it didn't really change in response to my interaction.
7(A) good, useful for reminding me the arriving emails, and add more fun for using computers. **at least it answers my questions and gives me some fun.**
8(A) It provides interesting company, especially with its **funny humour.** It made me laugh.
9(A) It showed its **friendly** character.

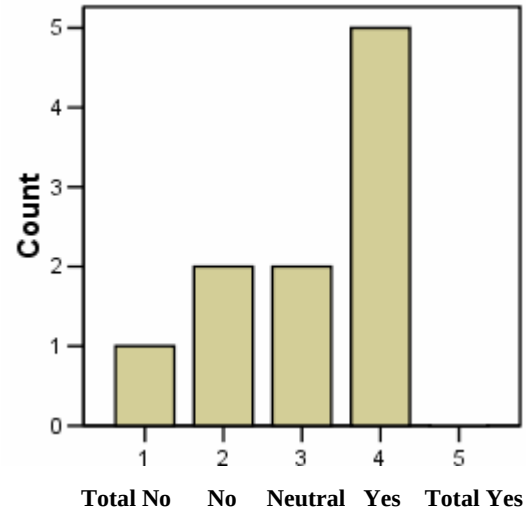
Quote 7.10: Comments of the Question "How was your satisfaction with the interaction with the Talking Head?"

7.4.3.3 Do you think the Talking Head is believable (believable means the Talking Head acts according to the expectations of the user)?

Figure 7.12(a) and Quote 7.11 show users' attitudes about the believability of the Email Agent. The distribution of Likert scales in Figure 7.12(a) shows 5 of 10 participants think it is believable. However, their qualitative comments show some of them do not really think it is believable. The reason may be that they realised the difficulties of making it believable: participant 3 said "I don't expect it to be truly life-like because I am aware of the difficulties in doing that!" but chose 4 for the Likert scale response to the question.

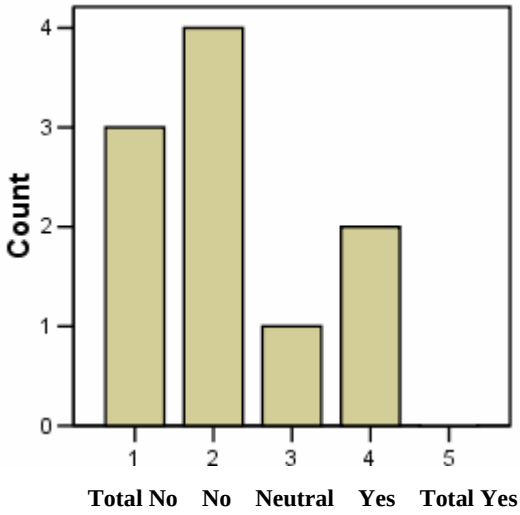
1(B) never talk too much with me
2(A) It simply isn't intelligent enough.
3(B) Difficult to make a judgement here - **I don't expect it to be truly life-like because I am aware of the difficulties in doing that!** On the other hand, the animations of the head were pretty natural, and this is where I would say it is 'believable'
4(B) Behaviour changes, at times counter to the purpose of the program
6(B) The only terms it understood were when I thanked it.
7(A) it could be designed like this.
8(A) I think it is the Talking Head's sense of humour.
10(A) It does seem to have a personality.

Quote 7.11: Comments of the Question "Do you think the Talking Head is believable (believable means the Talking Head acts according to the expectations of the user)?"



(a) Believable

Figure 7.12(a): Results of the Likert Scale for the Question "Do you think the Talking Head is believable (believable means the Talking Head acts according to the expectations of the user)?"



(b) Engaged

Figure 7.12(b): Results of the Likert Scale for the Question "Do you feel you engaged with the Talking Head (engaging means you find Talking Head entertaining and would like to spend time with)?"

7.4.3.4 Do you feel you engaged with the Talking Head (engaging means you find the Talking Head entertaining and would like to spend time with)?

Figure 7.12(b) and Quote 7.12 show users' attitudes of engagement towards the Email Agent. The distribution of Likert scale values in Figure 7.12(b) shows 7 of the 10 participants choose "no" or "total no" while only 2 participants chose "yes". Several reasons can be found from the qualitative comments:

1. Some users do not like the Email Agent's personality (see 1 in Quote 7.12).

2. Some users think the Email Agent is not intelligent and believable enough (see 2, 3 and 9 in Quote 7.12).
3. Some users did not engage with the Email Agent because of the scenario and they may engage with a personality-rich Email Agent when they play a game with it (see Section 7.4.3.5).

The above reasons indicate much more work needs to be done in the future and this future research will be discussed in Chapter 8. For example, the first reason indicates that the personality characteristics should be evaluated and filtered against the application domain and the characteristics such as “Hostile” and “Get upset easily” should have been filtered beforehand.

1(B) never talk too much with me
2(A) Again, it simply isn't intelligent enough.
3(B) It’s just a computer.
7(A) quite fun
8(A) It is funny.
9(A) Its brain needs more training.

Quote 7.12: Comments of the Question “Do you feel you engaged with the Talking Head (engaging means you find Talking Head entertaining and would like to spend time with)?”

7.4.3.5 Did you find the email application suitable for this use of an ECA with personality?

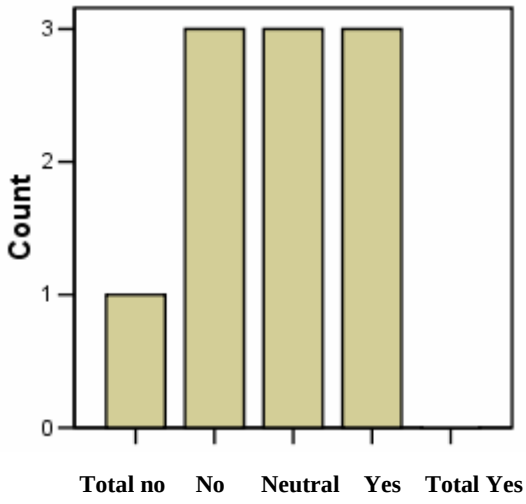


Figure 7.13: Results of the Likert Scale for the Question “Did you find the email application suitable for this use of an ECA with personality?”

2(A) (I had to look up ECA.) Existing email interfaces are more than sufficient. If there is a need for an audio alert for new mails, or for emails to be "read out" by the computer, surely this doesn't require a personality.
3(B) Again, since I only really care about the functionality, a mail notification beep would easily replace the function of the talking head in this app.
4(B) Distracting, and reading out the subject line in a work environment may not be a good idea without very good spam filters.
5(A) I don't. Perhaps elderly people or unskilled people might.
6(B) Correct rendering of pronunciation/names, although difficult obviously, would go a long way toward making it more of an aide.
7(A) remind me
8(A) This makes reading an email a much more enjoyable experience, especially with boring ones.
10(A) What is an ECA? Using the talking head for an email agent seems good. Analysing the content of emails could be useful. The talking head could give commentary/links for the content.

Quote 7.13: Comments of the Question "Did you find the email application suitable for this use of an ECA with personality?"

Figure 7.13 and Quote 7.13 show the users' attitudes to the scenario. Some of them did not think it was necessary for a personality-rich ECA to provide an email service either because they have a high level of computing experience, or the existing email applications are very efficient. This is reasonable since the Email Agent is just an experimental application to evaluate the personality model instead of a commercial product. Quote 7.14 shows their suggestion for other applications that will benefit the future research of designing an appropriate scenario for evaluating the ECA's personality.

2 Gaming is the most obvious one, I suppose, but it would really have to be integrated into the game. Perhaps it might be of some use in educational software for youngsters.
3 It's all about the target audience. Choose gadget-crazy people and I'm sure they'll love it. Other people (like me) aren't interested. It might even be that the target audience is all-important since I don't see the point of other talking head apps like AnaNova either.
5 For someone learning to use an application (any app) - to help explain certain processes.
7 mobile phone, car, home appliances etc
8 Online messaging and chatting services
9 Online joke narrater.
10 - rss feed notification - instant messaging, emails, rss, blogs, forums, etc all seem to be converging, so there might some ideas in that.

Quote 7.14: Comments of the Question "Please indicate other applications that may be suitable for an ECA with personality."

7.4.3.6 Are there any other comments you want to make?

Quote 7.15 shows qualitative comments from the question "Are there any other comments you want to make?". The comment of participant 3 indicated three problems of the Email Agent application that are good feedback for the study. The problems of the application will be discussed in Section 7.5.

3 Application specific: - **Doesn't display very large mails well.** You'll need to limit the size of the mail window to the desktop resolution and provide scroll bars. - **The Email agent spends most of its time minimised, or at least behind other windows.** This makes the animations almost superfluous. The new sidebars (MacOS and Windows Vista, I think) would be an ideal place to put the face alongside the clocks, etc. This would also have the advantage of using a smaller face (**the current face is too big**, and doesn't center properly if you shrink the application). It would also de-emphasize the lack of detail on the face.
7 the talking head is often stuck - get problem, I think it's the server's problem. If could overcome this problem, that would be much better.
8 I think the Talking Head is really entertaining and hope to see it maximise its potential in the future.

Quote 7.15: Comments of the Question "Are there any other comments you want to make."

7.5 Problems of Evaluation

The following lists the main problems discovered in the evaluation:

Sample Size is the biggest problem of the evaluation and this is discussed in Section 7.2.

Scenario is another problem that is discussed in Section 7.4.3.5.

Questionnaire: There are two significant problems with the questionnaire:

1. Statement Question "Like changes" is misunderstood by participants. They thought the statement was asked about "Like change their emotions" instead of "Like change their appearance" (see Section 7.3.2.3).
2. Some comments of participants are lead by the guide/descriptions for the questions (see Section 7.4.2.4 and Quote 7.7)

Participants are either computer science postgraduate students or computer science staff. This indicated a problem that such computing people may not share the views of non-computing people. One participant has implied this problem "Probably since I am a computing person … I am aware that there is a (large) target audience that doesn't

share my views." An ideal evaluation should involve both computing and none computing people and their attitudes need to be compared.

Dialogue Management: Some participants said the ECA does not understand what he/she said. This indicates two requirements:

1. Future work is needed to widen the Domain Knowledge of the ECA. Time is needed in order to provide enough domain knowledge for ECAs, and this is left as future work since this research is only one-year full time.
2. Participants need to be aware of the currently available topics before they use it.

Experimental Application: There are six problems or shortcomings of the experimental application indicated by participants:

1. Participant could not reply to email messages in the same interface (see Quote 7.3 participant 5).
2. The application should be able to analyse the content of emails and the ECA should be able to provide spoken commentary on, or link to, the content (see Quote7.13 participant 10).
3. The application does not display attachments (see Quote 7.3 participant 10).
4. The application does not display very large mails well (see Quote 7.15 participant 3).
5. The Email agent spends most of its time minimised, or at least behind other windows (see Quote 7.15 participant 3).
6. The face is initially set to be too big although the user can change the face size (see Quote 7.15 participant 3).

Statistical: Due to the limited sample size, detailed statistical tests have not been carried out.

1. Differences of Means from Group A to Group B were not carried out.
2. Detailed differences due to demographic and personality classification were not carried out.
3. Detailed changes over time were not carried out.

The above issues of the evaluation can also be considered as one of the achievements of this pilot study. Highlighting some of the problems is a useful caveat for future personality evaluation. It can be seen from this pilot study that in order to make future personality evaluations very successful, researchers need to find a large number of

participants to attend the evaluation for a long period of time. This is actually very difficult for Masters or even PHD level research.

However, the evaluation does have some positive results to support the hypotheses. Four obvious feelings - happy, angry, disgust and sad - were recognized by participants correctly. Participants correctly identified some of the personality characteristics such as “like to change its appearance”, “friendly”, “hostile”, “like moves”, “emotionally stable”. The Email Agent is believable is supported by 5 of 10 participants while it is denied by 3 participants.

7.6 Conclusion

This chapter detailed the data analysis and results of the evaluation.

- Section 7.2 described the demographic information of participants. The problem of the sample size was discussed in this section.
- Section 7.3 introduced the analysis of participants’ feelings and formative questions. Section 7.4 discussed the analysis of feedback of summative questions.
- Problems of the evaluation were discussed in Section 7.5.

The conclusions about each hypothesis will be presented in next chapter.

Chapter 8
Conclusions

8.1 Conclusions for Research Objectives

This research had a number of Objectives:

- Design a specification for an affective personality model based on existing research in the computer science and psychology literature.

Both *parent* theories and *research problem* theories were investigated (see Chapter 2) to find unexplored areas, or areas that needed further investigation. The *parent* theories came from the fields of Embodied Conversational Agents (ECAs) and personality theory while the *research problem* theories section investigated the application of a personality to ECAs. The specification of the affective personality model was designed as a psychological personality model by analysing the psychological description of typical personality and analysing the Scales and Items of International Personality Item Pool (IPIP) (Goldberg 2005) (see Chapter 4).
This objective was met.

- Convert the specification into a suitable computational model.

Section 4.3 discussed the issues of converting the psychological personality model into a computational personality model. The computational model, which included feelings, thoughts and behaviour modules, was developed and discussed in detail in the Section 4.3.
This objective was met.

- Implement the model within Curtin's ECA system.

Chapter 5 discussed the implementation of this personality model within Curtin's ECA system. An integrated architecture was implemented for this personality model. This personality model architecture formed the basis for all future personality research using the DM. An experimental application - Email Agent - was presented in this chapter, and the application fully implemented the personality model within Curtin's ECA system.
This objective was met.

- Evaluate the model in terms of user identification of the modelled personality, user satisfaction of the human computer interaction, and the model's ability to display emotion according to its personality.

Chapter 6 described the data collection of this research and Chapter 7 detailed the data analysis and results of the evaluation.
This objective was met.

8.2 Conclusions for Each Hypothesis

Hypothesis: A personality model will enhance the believability of an ECA and make human computer interaction more satisfying.

Hypothesis was partially supported by the results from the evaluations. Sections 7.4.3.2 and 7.4.3.3 analysed the users' feedback of satisfaction and believability. That the Email Agent is believable is supported by 5 of 10 participants while it is denied by 3 participants. Satisfaction of the interaction is denied by 4 participants who indicated that they were not satisfied with the interaction. Three possible reasons for this result concerning personality modelling, the scenario and the evaluation, and their affect on future development will be discussed in Section 8.3.

The personality model that was added to the Mentor System through hooks has made a significant contribution to the Mentor Dialogue Management System at Curtin. It is hoped that this will enhance the believability of future user dialogues and make the human computer interaction more satisfying for users.

Sub-hypothesis 1: An ECA equipped with a personality model can express emotion according to its personality characteristics.

Sub-hypothesis 1 was partially proven by the results of the data analysis. Six basic emotions were modeled as feelings of the ECA's, and were generated solely according to the personality characteristics of the ECA. Four obvious personality-generated emotions were recognized by participants correctly - happy, angry, disgust and sad - although the ECA's personality can control the expression of all six emotions. The other two emotions - surprised and afraid - were not noticed by participants since the chance of an Email Agent feeling surprised or afraid in this scenario was very low compared to feeling say happy or disgust.

Sub-hypothesis 2: A user can identify the personality or the personality characteristics of the ECA.

Sub-hypothesis 2 was partially proven by the results from the evaluations. Participants correctly identified most of the personality characteristics. These included ECA characteristics such as "like to change its appearance", "friendly", "hostile", "like moves", "emotionally stable".

However, the extrovert personality was identified as introverted since participants thought that "It didn't talk much." A design decision was made that the frequency of active responses was to be very low even for an extrovert ECA. This was to prevent it annoying participants and affecting their work, otherwise more participants may have dropped out before the two weeks.

Some participants recognized the friendly personality as unhappy since the ECA's physical appearance looks a bit unhappy and serious (see Figure 7.8).

8.3 Future Research

The design, implementation and evaluation phases of this research suggested that a useful system to add personality to an ECA has been developed. Nevertheless, more traits need to be added, and more applications need to be implemented to fully test the system. The evaluation is a pilot study giving valuable feedback on the questionnaire

and the methodology. Chapter 7 indicated some of problems detected during the evaluation. This section describes possible future work to solve these problems.

8.3.1 Personality Modelling

The number of personality characteristics in the computational personality model of this thesis is limited. Only the characteristics which are related to the traits of Friendliness, Extraversion and Neuroticism (FEN) have been analysed and modelled in the personality model (see Chapter 4). However many other personality characteristics are under research by psychologists, and some of them will be very useful for constructing a believable ECA. For example, an ECA with the personality characteristics “Cheer people up”, “Know how to captivate people” and “Start conversations” (Goldberg 2005). These may provide a satisfying interaction for a specific group of users. More characteristics are needed to be modeled in order to develop more believable personality-rich ECAs.

It is difficult to model all the possible personality characteristics. In future research, the usefulness of personality characteristics for ECAs needs to be evaluated against both the requirement of the user and the requirement of the application area. “Start conversations” may cheer up an open-minded user whilst making a utilitarian user upset. “Start conversations” will be a useful personality characteristic for an ECA who plays games with the user, while it will be useless for a personality characteristic of the Microsoft *Paper Clip*. Users of Microsoft Word most times do not want to start a conversation with the Microsoft *Paper Clip* except when they need help.

An indication of one reason why some participants were not satisfied with the Email Agent is that the personality characteristics have not been evaluated and filtered against the application domain before modelling. If this step had been done correctly, the characteristics such as “Hostile” and “Get upset easily” would have been filtered beforehand. The interaction would be more satisfying since only nice and useful characteristics would have been modelled. However, planning an experiment that only tests “nice” characteristics, does not answer questions about expected user-ECA clashes or confrontations from “not nice” characteristics.

8.3.2 Implementation

The personality model is limited to the existing techniques of ECA creation and manipulation since it needs to be implemented in an ECA system. Techniques introduced in Chapter 2 include MPEG-4 Facial Animation, Text to Emotional Speech (TTES) and Dialogue Management (DM). Future implementations of the personality model will depend upon the ECA framework. “Like to shake his leg” or “Like bow” cannot be implemented in the current agent as it only has a head instead of the whole body. “Cheer people up” cannot easily be implemented without very intelligent dialogue management system. Therefore, future work concerning the implementation of a personality model involves developing higher-level ECA techniques such as realising better physical models, and natural voice and animations.

8.3.3 Evaluation

- **Sample Size** is the biggest problem of the evaluation of this study (see Section 7.2). The problem of small sample size may occur in the future work of evaluation ECAs’ personality due to the difficulties of finding participants. The evaluation of other aspects of ECAs such as emotions and efficiency may need several minutes to several hours while evaluating personality needs several days to several weeks in order to perceive the personality characteristics correctly. Even psychologists do not know how long it should take to perceive a personality. Evaluating personality needs the participants’ collaboration for a long time and a participant may give up the evaluation if his personality mismatches the personality being interacted with. And yet it is necessary to conduct the experiment with the mismatch. The scenario of the personality evaluation experiment should be interesting and have useful functions in order to attract potential participants.
- **Participants** of this study are all computing people. There is a problem that computing people may not share the same views about an ECA’s personality as non-computing people. Future evaluations should involve various types of participants and their feedback can be grouped for analysis.

- **Scenarios**

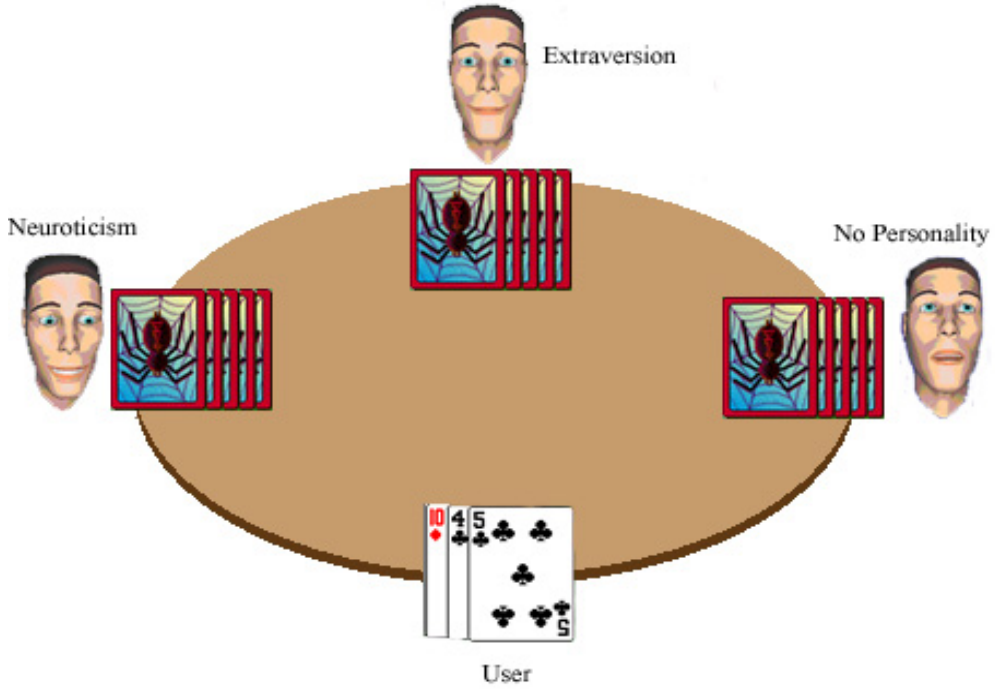


Figure 8.1: Card Game Scenario between User and Personality Rich ECAs

For evaluating personality rich ECAs in future research, better scenarios need to be designed. Games, education agents, joke narrating agents, news feed agents and sales agents may be considered. Figure 8.1 shows a card game scenario that was proposed by Curtin's Affective User Interface (AUI) group. The advantage of this scenario is whether the believability is enhanced and can be tested by a users' comparison of the interaction with a personality rich ECA vs. a personality neutral ECA. However, the duration of evaluation needs to be longer since a participant needs to perceive the personality characteristics of three ECAs at the same time.

Overall, the research has achieved its goals of design, implementation and evaluation of an affective personality model for an ECA. The personality model was proved to be partially affective since the ECA have six basic emotions and four obvious emotions, happy, angry, disgust and sad, were recognized by participants as generated by the personality model. More research needs to be done in the future, and future research will be difficult, due to the "human" nature of personality. Some phases of this study, especially the evaluation method and the evaluation results, should be beneficial to other researchers in this relatively new area of personality rich ECA construction.

References

Allport, GW 1937, *Personality: A psychological interpretation*, Henry Holt, New York.

André, E, Klesen, M, Gebhard, P, Allen, S & Rist, T 1999, 'Integrating Models of Personality and Emotions into Lifelike Characters', *Proceedings of the workshop on Affect in Interactions - Towards a new Generation of Interfaces in conjunction with the 3rd i3 Annual Conference*, eds. A Paiva & C Martinho, Siena, Italy.

André, E, Klesen, M, Gebhard, P, Allen, S & Rist, T 2000, 'Exploiting models of personality and emotions to control the behavior of animated interactive agents', *Proceedings of the workshop on "Achieving Human-Like Behavior in Interactive Animated Agents" in conjunction with the Fourth International Conference on Autonomous Agents*, ed. J Rickel, Barcelona, Catalonia, Spain.

Bandura, A 1977, *Social Learning Theory*, Prentice-Hall Inc., Englewood Cliffs, NJ.

Bandura, A 1986, *Social Foundations of Thought and Action: A Social Cognitive Theory*, Prentice Hall Inc., Englewood Cliffs, NJ.

Bandura, A 1989, 'Social Cognitive Theory. IN: Annals of Child Development', in R Vasta (ed.), vol. 6, Jai Press LTD., Greenwich, CT, pp. 1-60.

Barker, T 2003, *The Illusion of Life Revisited*, from http://www.soc.staffs.ac.uk/tb6/tim_barker_illusion_of_life_2.pdf

Bartneck, C 2002, 'Integrating the OCC Model of Emotions in Embodied Characters', *Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenges*, Melbourne.

Bates, J 1994, 'The role of emotion in believable agents', *Communications of the ACM*, vol. 37, no. 7, pp. 122-5.

Beard, S 2004, *MetaFace: A Virtual Face Framework and Metaphor*, PhD. Thesis, Curtin University of Technology.

Beard, S, Crossman, B, Cechner, P & Marriott, A 1999, 'FAQBot', *Pan Sydney Area Workshop on Visual Information Processing*, University of Sydney, Sydney.

Beard, S & Reid, D 2002, 'MetaFace and VHML: A First Implementation of the Virtual Human Markup Language', *AAMAS workshop on Embodied Conversational Agents - let's specify and evaluate them!*, AAMAS 2002, Bologna, Italy.

Beard, S, Reid, D & Shepherdson, R 2001, 'Believable and Interactive Talking Heads for Websites: Metaface and MPEG-4', *6th International Computer Science Conference: Active Media Technology*, Hong Kong.

Beizer, B 1990, *Software Testing Techniques*, Van Nostrand Reinhold, New York, NY.

Breese, J & Ball, G 1998, *Modeling Emotional State and Personality for Conversational Agents*, Microsoft Corporation, MSR-TR-98-41.

Bui, TD, Heylen, D, Poel, M & Nijholt, A 2002, 'ParleE: An adaptive plan-based event appraisal model of emotions', *Proceedings KI 2002, 25th German Conference on Artificial Intelligence*, eds. M Jarke, J Koehler & G Lakemeyer, Springer, Aachen, pp. 129-43.

Burns, AC & Bush, RF 2000, *Marketing Research*, third edition edn, Upper Saddle River, New Jersey.

Byrne, D, Griffitt, W & Stefaniak, D 1967, 'Attraction and Similarity of Personality Characteristics', *Journal of Personality and Social Psychology*, vol. 5, pp. 82-90.

Cassell, J, Sullivan, J, Prevost, S & Churchill, E (eds.) 2001, *Embodied Conversational Agents*, MIT Press, Cambridge, MA, USA.

Dam, HH & Souza, Sd 2002, 'Applying Talking Head Technology to a Web based Weather Service', *HF2002 Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenges*, Melbourne, Australia.

Dryer, DC 1999, 'Getting Personal with Computers: How to Design Personalities for Agents', *Applied Artificial Intelligence*, vol. 13, no. 3, pp. 273-95.

Eagly, AH, Ashmore, RD, Makhijani, MG & Longo, LC 1991, 'What is beautiful is good, but...:A meta-analytic review of research on the physical attractiveness stereotype', *Psychological Bulletin*, vol. 110, pp. 109-28.

Egges, A & Kshirsagar, S 2004, 'Generic Personality and Emotion Simulation for Conversational Agents', *Computer Animation and Virtual Worlds*, vol. 15, no. 1, pp. 1-13.

Ekman, P & Friesen, WV 1975, *Unmasking the Face. A guide to recognizing emotions from facial clues.*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Ekman, P, Friesen, WV & Ellsworth, P 1972, *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*, Pergamon Press Inc., New York.

Eptamedia 2005, *Eptamedia S.r.l.*, University of Genoa, Italy. Retrieved May 28, from http://www.eptamedia.com/faceanim/en-doc/en-faceanim.htm

Eysenck, SBG & Eysenck, HJ 1963, 'The validity of questionnaire and rating assessments of extraversion and neuroticism, and their factorial stability', *British Journal of Psychology*, vol. 54, no. 1, pp. 51-62.

Farlex 2005, *The Free Dictionary*, Huntingdon Valley, USA. Retrieved December 10, from http://www.thefreedictionary.com/feeling

Freud, S 1933, *New Introductory Lectures on Psycho-Analysis*, Standard Edition, James Strachey.

Ghasem-Aghaee, N & Ören, TI 2003, 'Towards Fuzzy Agents with Dynamic Personality for Human Behavior Simulation', *Proceedings of the 2003 Summer Computer Simulation Conference*, Montreal, PQ, Canada.

Goldberg, LR 2005, *International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences*, January 13, 2006. Retrieved August 28, from http://ipip.ori.org/

Gratch, J 2000, 'Émile: Marshalling Passions in Training and Education', *Proceedings of the 4th International Conference on Autonomous Agents*, Barcelona, Spain.

Gulland, EK 2001, *Psychological Modelling of an Asynchronous Idle-Time Personality for a Talking Head*, Honours Thesis Thesis, Curtin University of Technology.

Gustafson, J, Lindberg, N & Lundeberg, M 1999, 'The August Spoken Dialog System', in *Eurospeech'99*, Budapest, pp. 1151-4.

Hewstone, M, Fincham, FD & Foster, J (eds.) 2005, *Psychology*, Blackwell, Oxford.

Isbister, K 1995, 'Perceived Intelligence and the Design of Computer Characters', *Lifelike Computer Characters (LCC'95)*, Snowbird, Utah.

Isbister, K & Doyle, P 2002, 'Design and Evaluation of Embodied Conversational Agents: A Proposed Taxonomy', *AAMAS Workshop on Embodied conversational agents*, Bologna, Italy.

ISO/IEC 1998, *Text for ISO/IEC FDIS 14496-1 Systems. Online at http://www.cselt.it/mpeg/working_documents.htm*, ISO/IEC, 14496-1.

Kemmis, S & McTaggart, R 1988, in *The Action Research Planner, third edition*, Deakin University, Victoria, Australia.

Kshirsagar, S 2002, 'A Multilayer Personality Model', *Proceedings of 2nd International Symposium on Smart Graphics*, Hawthorne, New York, pp. 107-15.

Lavagetto, F & Pockaj, R 1999, 'The Facial Animation Engine: towards a high-level interface for the design of MPEG-4 compliant animated faces', *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, no. 2.

Leary, T 1957, *Interpersonal diagnosis of personality*, Ronald Press, New York.

Lester, JC & Stone, BA 1997, 'Increasing Believability in Animated Pedagogical Agents.' *First International Conference on Autonomous Agents*, ACM.

Lundeberg, M & Beskow, J 2005, *Developing a 3D-Agent for the August Dialog System*. Retrieved September 10, from http://www.speech.kth.se/multimodal/papers/mljb_avsp99.html

Marriott, A 2001c, 'A Java based Mentor System', in T Greening (ed.), *Java in the Computer Science Curriculum (to be published)*, LNCS, Springer.

Marriott, A 2002, 'A Facial Animation case study for HCI: the VHML-based Mentor System', in I Pandzic & R Forchheimer (eds), *MPEG-4 Facial Animation - The standard, implementations and applications*, John Wiley, London.

Marriott, A 2005, *A Software Based Mentor System*, PhD. Thesis, Curtin University of Technology.

Marriott, A to be published, 'A Java Based Mentor System', *Java in the Computer Science Curriculum*.

Marriott, A & Beard, S 2004, 'gUI. Specifying Complete User Interaction', in H Prendinger & M Ishizuka (eds), *Life-Like Characters. Tools, Affective Functions and Applications*, Springer-Verlag, Berlin, Germany, pp. 111-34.

Marriott, A, Beard, S, Haddad, H, Pockaj, R, Stallo, J, Huynh, Q & Tschirren, B 2001b, 'The Face of the Future', *Journal of Research and Practice in Information Technology*, vol. 32, no. 3/4.

Marriott, A, Beard, S, Stallo, J & Huynh, Q 2001d, 'VHML - Directing a Talking Head', *Active Media Technology Conference.*, Hong Kong.

Marriott, A, Pockaj, R & Parker, C 2001a, 'Internet Commerce and Software Agents: Cases, Technologies and Opportunities', in, Idea Group Publishing, United Kingdom, pp. 290–315.

Marriott, A, Pockaj, R & Parker, C 2001b, 'A Virtual Salesperson', in SM Rahman & RJ Bignall (eds), *Internet Commerce and Software Agents: Cases, Technologies and Opportunities*, Idea Group Publishing, pp. 290-315.

Marriott, A, Pockaj, R & Parker, C 2001a, 'The Face of E-commerce', in S Rahman & R Bignall (eds), *Internet Commerce and Software Agents – Cases, Technologies and Opportunities*, Idea Group Publishing.

Marriott, A & Shortland-Jones, B 2003, 'The Mentor System', *"Tutorial Dialogue Systems: With a View Towards the Classroom" Workshop Proceedings of 11th International Conference on Artificial Intelligence in Education*, ed. V Aleven, AIED2003, Sydney, Australia.

Marriott, A & Stallo, J 2002, 'VHML- Uncertainties and Problems... A discussion', *AAMAS Workshop on Embodied Conversational Agents - Let's Specify and Evaluate Them in conjunction with The First International Joint Conference on Autonomous Agents & Multi-Agent Systems*, Bologna, Italy.

McCrae, RR & John, OP 1992, 'An introduction to the five-factor model and its applications', *Journal of Personality*, vol. 60, pp. 175-215.

Melson, J 2000, *Giving Computers Personality*. Retrieved July 15, from http://www.honr.ttu.edu/GPHC/poster%20examples/NCHC.pdf

Moffat, D 1997, 'Personality Parameters and Programs', in *Creating Personalities for Synthetic Actors*, Springer Verlag, New York, pp. 120-65.

Myers, DG 2004, *Psychology*, Worth publishers, New York.

Nass, C, Isbister, K & Lee, E-J 2000, 'Truth Is Beauty: Researching Embodied Conversational Agents', in J Cassells, J Sullivan, S Prevost & E Churchill (eds), *Embodied Conversational Agents*, MIT Press, Cambridge, MA, pp. 374-402.

Nass, C, Moon, Y, Fogg, BJ, Reeves, B & Dryer, C 1995, 'Can Computer Personalities Be Human Personalities?' *International Journal of Human-Computer Studies*, vol. 43, no. 2, pp. 223-39.

O'Brien, R 1998, *An Overview of the Methodological Approach of Action Research*, from http://www.web.net/~robrien/papers/arfinal.html

Ochs, M, Niewiadomski, R, Pelachaud, C & Sadek, D 2005, 'Intelligent Expressions of Emotions', *The First International Conference of Affective Computing and Intelligent Interaction*, Beijing, China.

Ortony, A, Clore, G & Collins, A 1988, *The Cognitive Structure of Emotions*, Cambridge University Press, Cambridge.

Parke, FI & Waters, K 1996, *Computer Facial Animation*, A.K. Peters, Wellesley.

Patzer, GL 1985, *The physical attractiveness phenomena*, Plenum, New York.

Pelachaud, C & Bilvi, M 2003, ' Computational Model of Believable Conversational Agents', in M-P Huget (ed.), *Communication in Multiagent Systems: background, current trends and future*, Springer-Verlag, pp. 300-17.

Perry, C 1998, 'A Structured Approach for Presenting research Theses', *Australasian Marketing Journal*, vol. 6, no. 1, pp. 63-86.

Pervin, LA & John, OP 2001, *Personality: Theory and Research*, John Wiley & Sons, Inc., New York.

Picard, RW 1997, *Affective Computing*, M.I.T Press.

Picard, RW 1998, *Affective Computing*, The MIT Press, United States.

Popkins, NC 1998, *Natural characteristics that influence environment: How physical appearance affects personality*, Northwestern University. Retrieved July 12, from http://www.personalityresearch.org/papers/popkins2.html

Prendinger, H, Descamps, S & Ishizuka, M 2002, 'Scripting the Bodies and Minds of Life-like Characters', *The 7th Pacific Rim International Conference on Artificial Intelligence,*, Tokyo, Japan.

Reilly, WS 1996, *Believable Social and Emotional Agents*, Ph.D. Thesis Thesis, Carnegie Mellon University.

Reilly, WS & Bates, J 1992, *Building Emotional Agents*, School of Computer Science, Carnegie Mellon University, Pittsburgh.

Rousseau, D 1996, 'Personality in Computer Characters', *Proceedings of the 1996 AAAI Workshop on Entertainment and AI / A-Life*, Portland.

Rousseau, D & Hayes-Roth, B 1997a, *Interacting with Personality-Rich Characters, Tech. Report KSL-97- 06*, Stanford University.

Rousseau, D & Hayes-Roth, B 1997b, *A Social-Psychological Model for Synthetic Actors*, Knowledge Systems Laboratory, Stanford University.

Ruttkay, Z, Dormann, C & Noot, H 2002, 'Evaluating ECAs – What and how?' *AAMAS Workshop on Embodied conversational agents*, Bologna, Italy.

Ruttkay, Z & Noot, H 2005, 'Variations in Gesturing and Speech by GESTYLE', *International Journal of Human-Computer Studies Special Issue on 'Subtle Expressivity for Characters and Robots'*, vol. 62, no. 2, pp. 211-29.

Shepherdson, RH 2000, *The personality of a Talking Head*, Honours Thesis Thesis, Curtin University of Technology.

Skinner, BF 1953, *Science and Human Behavior*, Macmillan, New York.

Spiegel, MR 1961, *Theory and Problems of Statistics*, Scaum Publishing Co., New York.

Stallo, J 2000, *Simulating Emotional Speech for a Talking Head*, Honours thesis Thesis, Curtin University of Technology. http://www.vhml.org/theses/stalloj/thesis.doc.

Sun Microsystems, I 1998, *JavaMail API Design Specification Version 1.1*, Palo Alto, USA. Retrieved August 30, from http://www.orionserver.com/docs/specifications/JavaMail-1.1.pdf

Thompson, P 1989, *The Structure of Biological Theories*, State University of New York Press, Albany.

Vélasquez, JD 1997, 'Modeling emotions and other motivations in synthetic agent', *Proceedings of the AAAI Conference 1997*, AAAI Press and the MIT Press, pp. 10-5.

VHML 2001, *Virtual Human Markup Language*. Retrieved March 30, from http://www.vhml.org/

Wilkie, G 1993, *Object-Oriented Software Engineering - The Professional Developer's Guide*, Addison-Wesley, Sydney.

Wilson, I 1999, 'The Artificial Emotion Engine TM, Driving Emotional Behavior', *the AAAI 2000 Spring Symposium on Artificial Intelligence and Interactive Entertainment*, AAAI Press, pp. 76-80.

Wonisch, D & Cooper, G 2002, 'Interface Agents: Preferred appearance characteristics based upon context', *HF 2002 Human Factors Conference - tutorials and Workshops*, Melbourne, Australia.

Xiao, H, Reid, D, Marriott, A & Gulland, EK 2005, 'An Adaptive Personality Model for ECAs', *The First International Conference of Affective Computing and Intelligent Interaction (ACII2005)*, eds. J Tao, T Tan & RW Picard, Springer-Verlag Berlin Heidelberg, Beijing, China, pp. 637-45.

Appendices

A: Consent Form

The purpose of this questionnaire is to get valuable feedback from you. You will interact with the experiment application Email Agent. The Email Agent is the experiment application that was developed to implement the adaptive personality model. It can interact with the user in different manner based on its personality characteristics. It main function is to notify the user when the user receives a new mail. The feedback will be used to evaluate hypotheses of this research.

You will be required to evaluate the Email Agent and fill in the questionnaire. The questionnaire has 3 parts: You need to fill in the first part before using the program. This will take you half hour. The second part will be popped up several times when you use the program. This should take 5 minutes each time. Parts 3 will be filled in after the experiments and this part can take half hour. The program is invoked when you log in and should not be terminated except by logging out. It is hoped you will use the program for two to three weeks. If you wish to withdraw before two weeks, you can. Please inform us and we will uninstall the program. The participation is purely voluntary, you do not have to take part in this research and you are free to withdraw from it anytime.

Note:

- It is NOT you that is evaluated, it is the system.
- The Email Agent is a long-term interactive scenario, and that means you need to use it for at least a few weeks before answering the parts 3.
- The individual data collected will remain strictly confidential.
- The questionnaires are anonymous, as there is no need to record your name.
- User input will be recorded in a log file only identified by your nickname. Please use the same nickname during evaluation. A dialog box asking for name and password requires your user name not the nickname but neither the user name nor the password is logged.

If you have any questions, then feel free to ask them during the questionnaire or via email later.

I can be contacted: He Xiao (Adam) 12759585@cs.curtin.edu.au
(Or my supervisor): Andrew Marriott raytrace@cs.curtin.edu.au
Donald Reid donald@cs.curtin.edu.au

This study has been approved by the Curtin University Human Research Ethics Committee. Ethics Approval Number: ***MC 001/2006****. If needed, verification of approval can be obtained either by writing to the Curtin University Human Research Ethics Committee, c/- Office of Research and Development, Curtin University of Technology, GPO Box U1987, Perth, 6845 or by telephoning 9266 2784.*

Consent Form:

- I have been informed of and understand the purpose of the study.
- I have been given an opportunity to ask questions.
- I agree to participate in the study as outlined to me.

Name: _____ Signature: _____

Date: _____

Questionnaire:

http://www.aui.computing.edu.au/projects/Personality/Questionnaire1.htm

B: Demographic Questionnaire (Online)

1 Your nickname

2 Your Gender

female male

3 Age (Click for choices)

4 Is English your first language?

yes no

If no, what is your first language?

5 What is your primary cultural or racial identification? (Click for choices)

6 How long have you used email? years

7 Have you ever used ICQ, MSN or similar system before?

yes no

If yes, what have you used, and how long have you used?

ICQ:

MSN:

Others:

8 Have you ever used a Talking Head system before?

yes no

If yes, how long have you used? months

Comments:

9 Have you ever used a dialogue management system such as Mentor system before?

yes no

If yes, what have you used, and how long have you used? ☐ months
Comments:

10 Please use the rating scale below to describe how accurately each statement describes *you*. Describe yourself as you generally are now, not as you wish to be in the future. Please read each statement carefully, and then fill in the bubble that corresponds to the number on the scale.
(Permitted by International Personality Item Pool http://ipip.ori.org/)

1 Very Inaccurate
2 Moderately Inaccurate
3 Neither Inaccurate nor Accurate
4 Accurate
5 Very Accurate

10.1 Am the life of the party.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.2 Feel little concern for others.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.3 Am always prepared.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.4 Get stressed out easily.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.5 Have a rich vocabulary.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.6 Don't talk a lot.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.7 Am interested in people.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.8 Leave my belongings around.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.9 Am relaxed most of the time.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.10 Have difficulty understanding abstract ideas.

Very Inaccurate	1 ○	2 ○	3 ○	4 ○	5 ○	Very Accurate

10.11 Feel comfortable around people.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.12 Insult people.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.13 Pay attention to details.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.14 Worry about things.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.15 Have a vivid imagination.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.16 Keep in the background.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.17 Sympathize with others' feelings.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.18 Make a mess of things.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.19 Seldom feel blue.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.20 Am not interested in abstract ideas.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.21 Start conversations.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.22 Am not interested in other people's problems.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.23 Get chores done right away.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.24 Am easily disturbed.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.25 Have excellent ideas.

Very Inaccurate	1	2	3	4	5	Very Accurate

10.26 Have little to say.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.27 Have a soft heart.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.28 Often forget to put things back in their proper place.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.29 Get upset easily.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.30 Do not have a good imagination.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.31 Talk to a lot of different people at parties.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.32 Am not really interested in others.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.33 Like order.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.34 Change my mood a lot.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.35 Am quick to understand things.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.36 Don't like to draw attention to myself.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.37 Take time out for others.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.38 Shirk my duties.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.39 Have frequent mood swings.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.40 Use difficult words.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.41 Don't mind being the center of attention.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.42 Feel others' emotions.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.43 Follow a schedule.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.44 Get irritated easily.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ◉	5 ☐	Very Accurate

10.45 Spend time reflecting on things.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.46 Am quiet around strangers.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.47 Make people feel at ease.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.48 Am exacting in my work.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.49 Often feel blue.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

10.50 Am full of ideas.

Very Inaccurate	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	Very Accurate

C: Formative Questionnaire

D: Summative Questionnaire (Online)

Your nickname

Part A

1 How long did you use the Email Agent system? days
Comments:

2 Please describe the characteristics showed by the Talking Head concerning his/her <u>appearances</u>
(for example you may answer "It always makes itself look cool").

3 Please describe the characteristics showed by the Talking Head concerning his/her <u>behaviors</u>
(for example you may answer "It always talks a lot and keep moving").

4 Please describe the characteristics showed by the Talking Head concerning his/her <u>emotions</u>
(happy, angry, disgust, sad, surprise etc).

5 Please describe the characteristics showed by the Talking Head concerning his/her <u>thoughts</u>
(self-control, optimism or pessimism etc).

6 Please use the rating scale below to describe how accurately each personality labels describes the *Talking Head*.
(If you don't understand the meaning of the word fully please choose *Not Available*.)

6.1 Aggressive

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

6.2 Calm

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

6.3 Emotionally stable

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

6.4 Extroverted

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

6.5 Friendly

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

6.6 Hostile

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

6.7 Introverted

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

6.8 Neurotic

Very Inaccurate	Inaccurate	Neutral	Accurate	Very Accurate	Not Available
○	○	○	○	○	○

Comments:

7 Please describe the personality of the Talking Head in your own words.

Part B

8 Do you agree that a Talking Head should have personality?

Strongly disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly agree ☐

Comments:

Why do you feel that a Talking Head needs or does not need a personality?

9 How was your satisfaction with the interaction with the Talking Head?

Total unsatisfied ☐ Unsatisfied ☐ Neutral ☐ Satisfied ☐ Total satisfied ☐

Why?

If you were unsatisfied with the talking head, would you be satisfied by the talking head with other personality characteristics?

☐ yes ☐ no

What kind of Talking Head would you prefer (for example you may answer "I prefer an optimistic and joyful talking head")?

Comments:

10 Do you think the Talking Head is believable (believable means the Talking Head acts according to the expectations of the user)?

Total no ☐ No ☐ Neutral ☐ Yes ☐ Total yes ☐

Why?

11 Do you feel you engaged with the Talking Head (engaging means you find Talking Head entertaining and would like to spend time with)?

Total no ☐ No ☐ Neutral ☐ Yes ☐ Total yes ☐

Why?

12 Did you find the email application suitable for this use of an ECA with personality.

Total no ☐ No ☐ Neutral ☐ Yes ☐ Total yes ☐

Why?

Please indicate other applications that may be suitable for an ECA with personality.

13 Are there any other comments you want to make.

E: IPIP Scales and Items

NEUROTICISM (NEO Personality Inventory)

10-item scale

+ keyed	Often feel blue.
	Dislike myself.
	Am often down in the dumps.
	Have frequent mood swings.
	Panic easily.
– keyed	Rarely get irritated.
	Seldom feel blue.
	Feel comfortable with myself.
	Am not easily bothered by things.
	Am very pleased with myself.

20-item scale (Alpha = .91)

+ keyed	Often feel blue.
	Dislike myself.
	Am often down in the dumps.
	Have frequent mood swings.
	Panic easily.
	Am filled with doubts about things.
	Feel threatened easily.
	Get stressed out easily.
	Fear for the worst.
	Worry about things.
– keyed	Seldom feel blue.
	Feel comfortable with myself.
	Rarely get irritated.
	Am not easily bothered by things.
	Am very pleased with myself.
	Am relaxed most of the time.
	Seldom get mad.
	Am not easily frustrated.
	Remain calm under pressure.
	Rarely lose my composure.

EMOTIONAL STABILITY (Big-Five Factor)

10-item scale

+ keyed	Am relaxed most of the time.
	Seldom feel blue.
– keyed	Get stressed out easily.
	Worry about things.
	Am easily disturbed.
	Get upset easily.
	Change my mood a lot.
	Have frequent mood swings.
	Get irritated easily.
	Often feel blue.

20-item scale

+ keyed	Am relaxed most of the time.
	Seldom feel blue.
	Am not easily bothered by things.
	Rarely get irritated.
	Seldom get mad.
– keyed	Get stressed out easily.
	Worry about things.
	Am easily disturbed.
	Get upset easily.
	Change my mood a lot.
	Have frequent mood swings.
	Get irritated easily.
	Often feel blue.
	Get angry easily.
	Panic easily.
	Feel threatened easily.
	Get overwhelmed by emotions.
	Take offense easily.
	Get caught up in my problems.
	Grumble about things.

INTROVERSION (Gough's California Psychological Inventory)

+ keyed	Don't like to draw attention to myself.
	Keep in the background.
	Dislike being the center of attention.
	Don't talk a lot.
– keyed	Don't mind being the center of attention.
	Take charge.
	Want to be in charge.
	Am the life of the party.
	Can talk others into doing things.
	Seek to influence others.

INTROVERSION (Cattell's 16 Personality Factors)

+ keyed	Want to be left alone.
	Prefer to do things by myself.
	Enjoy spending time by myself.
	Seek quiet.
	Don't mind eating alone.
	Enjoy silence.
	Enjoy my privacy.
– keyed	Enjoy being part of a group.
	Enjoy teamwork.
	Can't do without the company of others.

EXTRAVERSION (NEO Personality Inventory)

20-item scale

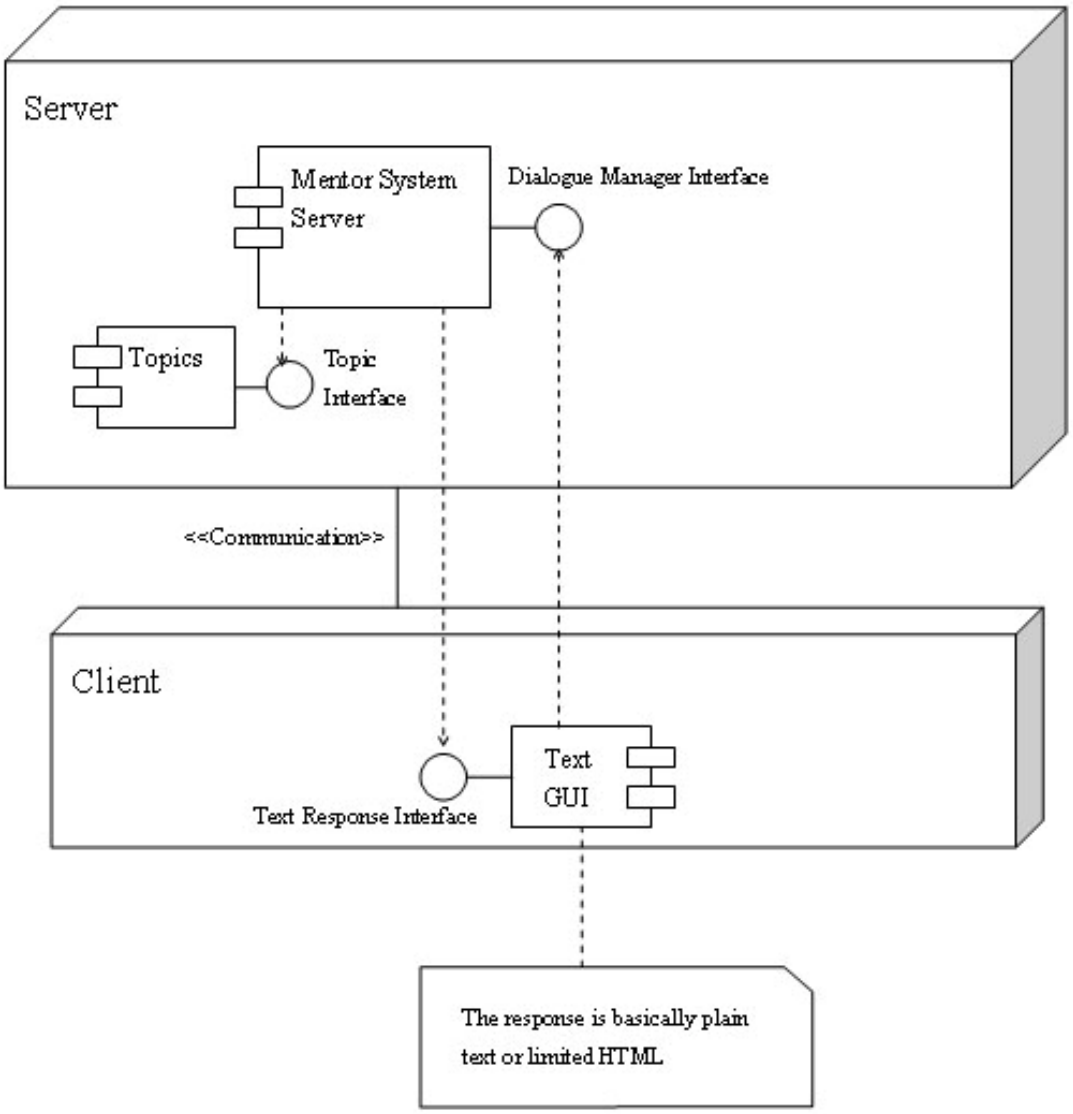
+ keyed	Feel comfortable around people.
	Make friends easily.
	Am skilled in handling social situations.
	Am the life of the party.
	Know how to captivate people.
	Start conversations.
	Warm up quickly to others.
	Talk to a lot of different people at parties.
	Don't mind being the center of attention.
	Cheer people up.
– keyed	Have little to say.
	Keep in the background.
	Would describe my experiences as somewhat dull.
	Don't like to draw attention to myself.
	Don't talk a lot.
	Avoid contacts with others.
	Am hard to get to know.
	Retreat from others.
	Find it difficult to approach others.
	Keep others at a distance.

FRIENDLINESS (NEO Personality Inventory)

+ keyed	Make friends easily.
	Warm up quickly to others.
	Feel comfortable around people.
	Act comfortably with others.
	Cheer people up.
– keyed	Am hard to get to know.
	Often feel uncomfortable around others.
	Avoid contacts with others.
	Am not really interested in others.
	Keep others at a distance.

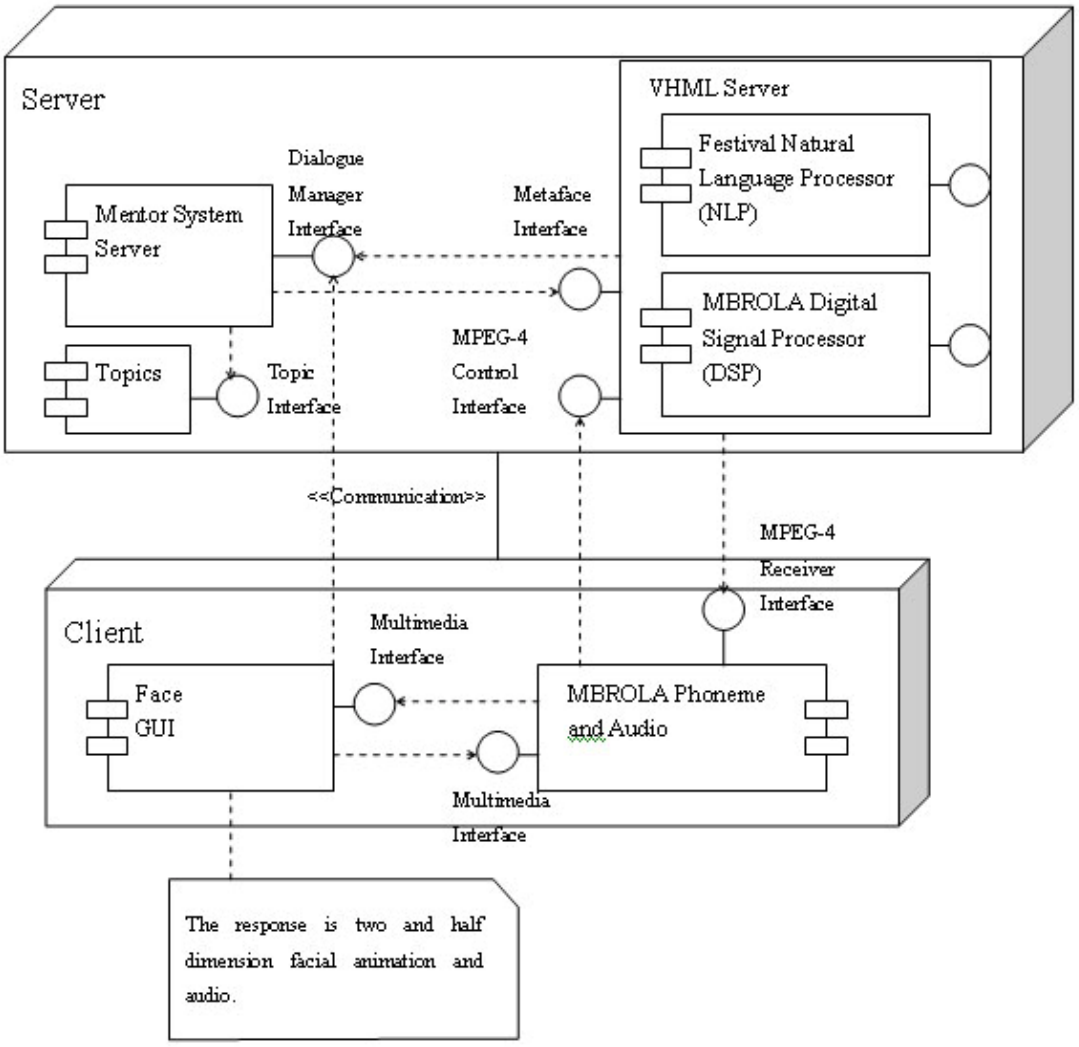
F: UML Deployment Diagrams

Phase One - Mentor System (Marriott 2005)



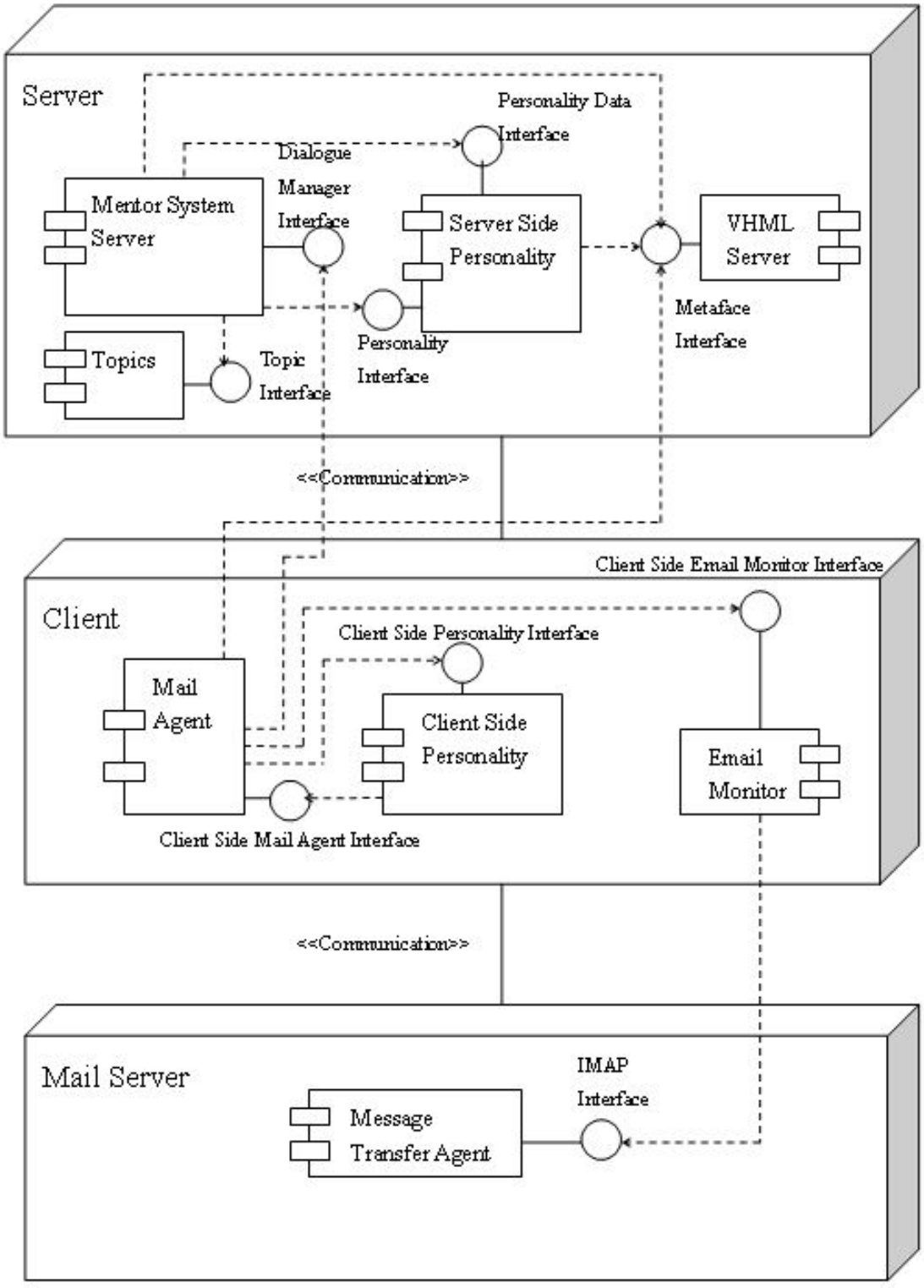
This shows the deployment diagram of the existing Mentor System used in this research.

Phase Two - Metaface System (Beard 2004)



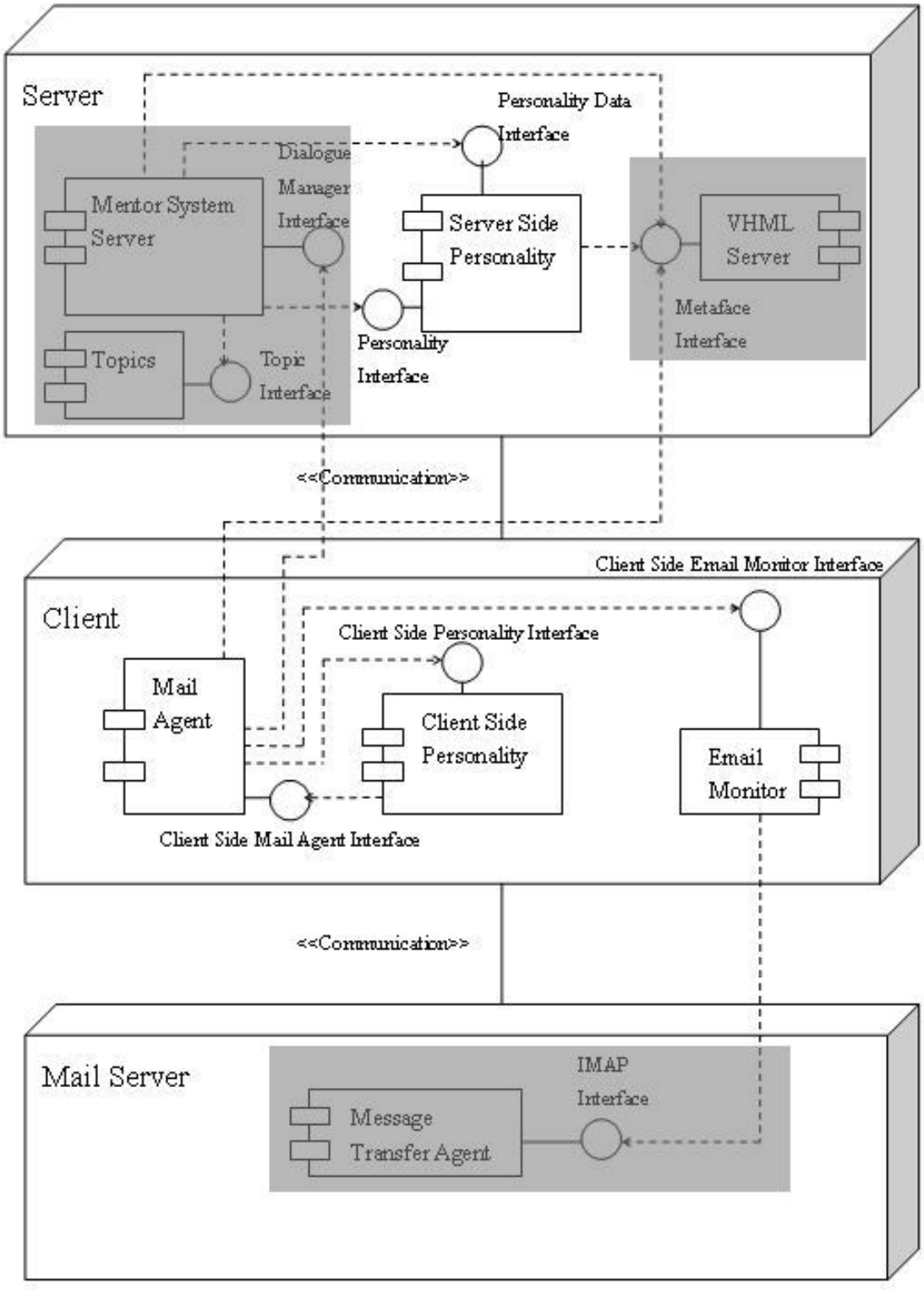
This shows the deployment diagram of the existing Metaface Framework used in this research.

Phase Three - Email Agent System



This shows the complete deployment diagram of the designed and implemented Email Agent System.

Implemented components of the Email Agent System:



This show the components designed and implemented for this research and how they integrate into the entire system. Greyed areas are from previous research.

The following tables show the details of interfaces which implemented in the Email Agent System.

<<interface>> PersonalityInterface
Void <init>() Void <init>(String) Void run() Void run_server() Void run_client() Void init() Void setDialogueManagerInterface(DialogueManagerInterface) Boolean initialisePersonality (DialogueManagerInterface, PersonalityDataInterface) Response beforeStateResponse (Context, Response, PersonalityDataInterface) Response beforeCheckFollowupResponse (Context, Response, PersonalityDataInterface, TopicsInterface.Vector) Response afterCheckFollowupResponse (Context, Response, PersonalityDataInterface, TopicsInterface.Vector) Response beforeStateLessResponse (Context, Response, PersonalityDataInterface) Response beforeCheckArea (Context, Response, PersonalityDataInterface, TopicInterface) Response beforeCheckTopic (Context, Response, PersonalityDataInterface, TopicInterface) Response afterCheckTopic (Context, Response, PersonalityDataInterface, TopicInterface) Response afterAllTopic (Context, Response, PersonalityDataInterface) Response finalResponseIsNull (Context, Response, PersonalityDataInterface) Response unknownTopicChosen (Context, Response, PersonalityDataInterface) Response responseGeneration(Context, Response, PersonalityData) Response lastWord(Context, Response, PersonalityDataInterface) Void tidyUp(Context, Response, PersonalityDataInterface) Void notifyDialogueMangerListener(DialogueManagerInterface, String, Object, Object) Void initialGreeting(DialogueManagerInterface, Response, PersonalityDataInterface) Void goodBye(Context, Response, PersonalityDataInterface) PersonalityDataInterface loadPersonalityData(ObjectInputStream, PersonalityDataInterface) PersonalityDataInterface savePersonality(ObjectOutputStream, PersonalityDataInterface) Boolean initialiseTopic(DialogueManagerInterface) Response checkTopic(Context) Int getTopicsActiveQueryWeight(DialogueManagerInterface) Void topicsActiveQueryIntialise(DialogueManagerInterface) Response getTopicsActiveQueryResponse(DialogueManagerInterface) Response processActiveQuery(DialogueManagerInterface, Response)

<<interface>>

PersonalityDataInterface

Void <init>()
Void init()
Void init(String)
String javaDirectory()
Int getValueFromPersonality(String, Int, Int)
Void createGUI()
Void windowClosed(WindowEvent)
Void windowClosing(WindowEvent)
Void windowDeactivated (WindowEvent)
Void windowDeiconified (WindowEvent)
Void windowIconified (WindowEvent)
Void windowOpened (WindowEvent)
Void windowActivated (WindowEvent)
Void actionPerformed (ActionEvent)
Void friendDialog ()
Void loginDialog()
Void nicknameDialog()
Void emotionSlidersDialog()
Void questionnaireDialog()
Void run()
Void start()
DialogueManagerStimulusInterface instantiateStimulusClass(String)
Object[] getDataObjects()
Void processCommand(Int, InputStream)
Void sendDataToTargets(Int, InputStream)
Void sendStringToServer (String, String)
Void sendStringToEvaluation (String, String)
Void processMark(String[])
Void processPerson(String[])
Dimension getPreferredSize()
String to String()
Void <clinit>()

<<interface>>

ClientSidePersonalityInterface

Void <init>()
Void init()
Void init(String)
String javaDirectory()
Int getValueFromPersonality(String, Int, Int)
Void createGUI()
Void friendDialog ()
Void loginDialog()
Void nicknameDialog()
Void emotionSlidersDialog()
Void questionnaireDialog()
Void run()
Void start()
DialogueManagerStimulusInterface instantiateStimulusClass(String)
Object[] getDataObjects()
Void processCommand(Int, InputStream)
Void sendDataToTargets(Int, InputStream)
Void sendStringToServer (String, String)
Void sendStringToEvaluation (String, String)
Void processMark(String[])
Void processPerson(String[])
Dimension getPreferredSize()
String to String()

<<interface>>

ClientSideMailAgentInterface

Void init(String)
Dimension getPreferredSize()
Void windowClosed(WindowEvent)
Void windowClosing(WindowEvent)
Void windowDeactivated (WindowEvent)
Void windowDeiconified (WindowEvent)
Void windowIconified (WindowEvent)
Void windowOpened (WindowEvent)
Void windowActivated (WindowEvent)
Void actionPerformed (ActionEvent)
Void run()
Void start()
String to String()

<<interface>>

ClientSideEmailMonitorInterface

Void <init>()
Void init()
Void init(String)
Object[] getDataObjects()
Void processCommand(Int, InputStream)
Void sendDataToTargets(Int, InputStream)
Void sendStringToServer (String, String)
Void sendStringToEvaluation (String, String)
Void processMark(String[])
DialogueManagerStimulusInterface instantiateStimulusClass(String)
Void sendDataToTargets(Int, InputStream)
Void sendStringToServer (String, String)
Void sendStringToEvaluation (String, String)
Void processMark(String[])
Void processPerson(String[])
Dimension getPreferredSize()

G: CDROM

To accompany this thesis is a CDROM, containing the documents, source-code and a live movie. As with any live recordings this suffers from ambient noise, but indicates some early work exhibited at the First International Conference of Affective Computing and Intelligent Interaction (ACII2005).