CREATING AUTOMATED VIRTUAL HUMANS

by

Raghavi Sakpal

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Computing and Information Systems

Charlotte

2015

Approved by:

____
Dr. Dale-Marie Wilson

____
Dr. Zachary Wartell

____
Dr. Christine Elnitsky

____
Dr. Wlodek Zadrozny

____
Dr. Celine Latulipe

ABSTRACT

RAGHAVI SAKPAL. Creating automated virtual humans. (Under the direction of DR. DALE-MARIE WILSON)

Virtual Humans (VHs) are highly efficient and effective task oriented tools for various social and collaborative environments. VHs have the ability to replicate human like verbal (speech) and non-verbal (gestures, facial expressions) interactive behaviors. They are presently being used in training, testing, improving communication skills and practice, with the added advantages of; fidelity of presentation; ability to portray a much wider range of personality; appearance; emotions; behavior; and the potential to be available anywhere, anytime, and at a low cost.

My research objective is to develop intelligent VHs with the ability to portray emotions and generate behaviors based on their history, education, personal experiences and cognitive state of mind. Emotions described as a state of feeling in the sense of an affect are often intertwined with mood, temperament, personality, disposition, and motivation. Over the years various agent architectures have been developed using computational models of emotions to drive the behavior of VHs. A common deficit seen across these agent architectures is their inability to incorporate one's personal experiences, history, and education while computing the decision making process for behavior generation. To this effect an agent architecture, called Culturally Modified Agent Architecture (CMAA), was developed with the ability to generate autonomous VHs whose cognitive state of mind is driven by three main factors: 1) agent's belief (past history and personal experiences), 2) agent's personality and the 3) mood of the

agent. The two main components driving the CMAA architecture are- the appraisal model (responsible for appraising the events based on agent's belief) and the emotion model (responsible for generating a set of emotions based on appraisal variables and PAD (Pleasure-Arousal-Dominance) rules of personality and emotions).

To test the feasibility of developing autonomous intelligent VHs using the proposed CMAA agent architecture, a VH prototype was implemented within the clinical setting. A VH prototype termed as VSP (Virtual Standardized Patient) portraying an OEF (Operation Enduring Freedom)/OIF (Operation Iraqi Freedom)/OND (Operation New Dawn) Veteran, exhibiting symptoms associated with mild TBI (Traumatic Brain Injury), was developed as a screening tool for evaluation purposes. Currently two different versions of the VSP exist- Version 1, where the VSPs behavior and emotions are scripted by the experts based on observations of a typical screening of mild TBI patient, and Version 2, where the behavior and emotions are automated and driven by the CMAA agent architecture. Evaluation studies were designed to 1) test the validity and believability of the VSP portraying symptoms of mild TBI and 2) test the effectiveness of using the VSP as a training tool to practice diagnostic evaluation and improve communication between a patient and a provider within a clinical setting.

This dissertation presents an in-depth review of the development and social impact of VHs across various domains. It describes the working of the CMAA agent architecture and its structural components for creating personalized, autonomous, intelligent agents. The dissertation then describes the design and development of the VSP as a diagnostic training tool for evaluation of mild TBI (Traumatic Brain Injury). Lastly an in-depth description of the evaluation studies along with the results are presented. The dissertation concludes by highlighting the potential possibilities of future work.

ACKNOWLEDGMENTS

I would like to gratefully acknowledge the immense support and guidance of my advisor and mentor Dr. Dale-Marie Wilson. I would also like to thank my committee members Dr. Christine Elnitsky, Dr. Zachary Wartell, Dr. Wlodek Zadrozny, and Dr. Celine Latulipe for their advice. I would like to acknowledge the members and friends of the Video & Image Analysis Lab in the Department of Computer Science, notably Kieran Jordine, for their support and technical assistance in my research. I would like to thank Dr. Katherine Taber from the Veteran Affairs for her help and guidance on the research project. I would also like to acknowledge and thank the faculty, and staff of the UNC-Charlotte community for guiding and supporting me in my research and career choices. Lastly, I would like to thank the students of UNC-Charlotte for enriching me with their experiences and making me a better teacher.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Background

The development of Intelligent Virtual Humans (VHs) has been a subject of exponential growth due to their accessibility, usability and adaptability in many application areas such as education, training, human-computer interfaces and entertainment [33] [35] [52]. These VHs that are modeled after humans often have the ability to portray common human-like morbidities, such as gestures, facial expressions, emotions, personalities, and speech, which are "transparent" to the user [33]. Research has shown that when users identify with these common morbidities of human behavior and attitude towards VHs, they tend to relate, learn and perform tasks more effectively [52] [35] [33] [36].

Nass proposes that humans tend to interact with computers as they do with other humans, based on the social rules and stereotypes that govern human interaction [33] [52]. Nass suggests that these social rules that govern human-to-human interaction may also apply to human to computer interactions. He also suggests that factors like verbal output (speech), appearance, behavior and personality of a VH can play an important role in social or collaborative environments, like education, medicine, games, military, and virtual reality [33]. However, some of these factors, such as appearance, personality, and behaviors that are necessary to facilitate effective human to VH communication are based on introspection rather than careful consideration of the tasks and users involved within the interactive system [33].

Evidential research has demonstrated that advances in VH interfaces are effective in supporting multi-modal interaction with users for a variety of tasks [33] [35] [52]. Over the past years, VHs have been used in socially diverse collaborative task-oriented environments such as virtual tutors like Dr. Chestr [26] to improve learning outcomes in Computer Science education; virtual trainers for mission rehearsal exercises [35] to facilitate servicemen in effective military training; and virtual task-oriented companions in social settings like STEVE [53] for training in engineering. The use of VHs, in the health care industry, has also increased due to the increasing demands on the education and training of health care professionals. These VHs with their ability to 'don' different virtual formats (from physical simulators to artificial patients) provide health care professionals an opportunity to practice communication, practice making diagnostic and therapeutic decisions, and practice psychotherapy training in a safe working and interactive environment [33] [36] [37] [80] [81] [82]. My work focuses on this area of research, i.e. creating an interactive system using VHs to standardize the diagnostic process of training and improve communication within a clinical setting.

1.2 VH Development

The goals for developing VHs have evolved throughout the years with the introduction of innovative technologies and resources available [12]. But the focus of VH development has always been across three main disciplines:

- Appearance Modeling: This means that visually the VHs should look like real people, that includes face and body shapes of real human, skin textures and realistic clothes.

- Realistic, smooth and flexible motion in any situation: Make the VHs to portray natural motions and animation control.

- Realistic, High-level behavior: Responsible for the cognitive behavior of the VH. Involves computational models of A.I. and agent technology.

Nowadays, a lot of interest from both industry and research exists for a VHs and their virtual environments. A lot of new techniques are being developed to improve the simulation in general, to add more visual details and make the interactions between humans and VHs more natural and believable [2][3]. Believability is a critical aspect of a VH which can be augmented by surface realization and intelligent behavior. Although advances in computer graphics has led to realistic surface realization, yet the VHs behavior is monotonous due to its scripted nature. To develop the behavior of VHs, traditional agent architecture follows a dualist perspective which decomposes the agent into a mind and a body. Mind as an abstract layer provides the agent with cognitive functionalities. It receives perceptions from the body, makes decisions, and sends the decisions in terms of abstract actions to the body. The body as an embodied layer animates the received actions within the virtual environment and provides the mind with perceptions acquired from its virtual sensors.

Some other approaches for VH development include categorizing the agents into logic-based, reactive, belief-desire-intention (BDI) layers of reasoning[23]. The BDI agent architecture has been widely utilized as a VH architecture. Logic-based agents tend to exploit symbolic logic deductions and cannot handle uncertainties, whereas cognitive agents such as BDI provide deliberative decision making capabilities for long temporal horizons. However, they cannot react to the situations which need instant responds. Furthermore, knowledge representation is a main challenge in these architectures. Reactive agents (i.e. behavior-based agents in robotics literature) couple the control and decision making mechanisms to the current local sensory information to provide real-time reactions. Although this approach minimizes the complexity of the representational structures and provides quick responses to dynamic environments, it is not scalable and suffers from the lack of reasoning capabilities and task-oriented behaviors. To develop a truly believable intelligent VH, an agent architecture should be able to incorporate cognitive characteristics of recognition, decision making, perception, situation assessment, prediction, problem solving, planning, reasoning, belief maintenance, execution, interaction and communication, reflection, and learning [1].

1.3 Research Problem and Questions

As highlighted in the earlier section, several challenges exist in the development of intelligent VHs and their interactive communication system. The existing technology to develop and evaluate VHs is individualistic and not yet standardized [54]. These individualistic approaches span to several different research areas that contribute towards the architecture of VH development including animation and rendering of virtual characters, planning and discourse modeling of their behavior, unobtrusive forms of human identification, and real-time speech recognition and synthesis [1][51]. These architectural models tend to be more focused towards the goals (tasks that need to achieved based on the environment) and actions (communication or gesture related actions that need to be performed to realize the goals) of the VH development rather than answering the following questions 1) "What drives the behavior and decision making of the VHs?", 2) "Does the VHs past experiences and personal beliefs determine their goals and actions?" and 3) "Do these personal experiences and beliefs shape the verbal and nonverbal behavior of the VH within an interactive environment?"

To address these issues and investigate the use of VHs as effective training tools in verbal and non-verbal communication, I propose the following research questions:

1. Is it possible to create autonomous VHs whose behavior is determined or affected by one's personal history, education, family background, work history and personal experiences?

2. Is it possible to use VHs as a training tools to improve VH-to-user communication, based on its verbal and non-verbal behavior?

1.4 Dissertation Overview

This dissertation provides an in-depth description of the background, research problem and hypothesis, implementation details of the research project and present results from the evaluation study conducted to validate the implementation. Chapter 1 focused on the summary of the overall dissertation project, while highlighting the research problems an questions raised. In Chapter 2 a more detailed literature survey is provided leading towards the problem definition of this dissertation. Chapter 3 highlights the research questions and presents a preview of my research contributions.

Chapter 4 describes the preliminary work done in the development of the first basic agent architecture incorporating personality and natural language dialog. To test the feasibility of developing a VH with personality, Dr. Chestr a virtual game-show host was designed and implemented within an educational setting as a virtual tutor quizzing students on the basics of C++. Implementation details of Dr. Chestr and results from the evaluation studies are also presented within Chapter 4. Chapter 5 describes the implementation details of our current agent architecture CMAA for developing autonomous intelligent agents with the ability to make decisions and portray behavior based on personality, mood and cultural beliefs. In this chapter, each component of the CMAA architecture is explained in detail and mapping rules are presented across the various components. To test the feasibility of CMAA, a virtual standardized patient (VSP) portraying a war veteran and exhibiting the symptoms of mild TBI (Traumatic Brain Injury) was developed as a diagnostic tool. Chapter 6 presents the design and development of two different versions of VSP interface. Finally in Chapter 7, design and description of the evaluation study is presented along with the hypothesis and the results from the user study. Chapter 7 also provides details and results from the heuristic evaluation conducted by the experts to test the physical simulation of the VSP. The evaluation questionnaire from all the evaluation studies (Dr. Chestr, VSP with mild TBI) have been described within the Appendix. The dissertation concludes by highlighting each research problem and its results in Chapter 8. Chapter 8 also provides an insight on future possibilities of this research project.

CHAPTER 2: LITERATURE SURVEY

This chapter describes the various aspects of Virtual Human (VH) Development, starting with their Social Effects and Applications and going onto the various development architectures and process involved in creating effective and believable VHs and their interactions within a task-based environment.

2.1 Virtual Humans/Embodied Conversational Agents

VHs, also known as Embodied Conversation Agents, are computer generated models of people who are capable of carrying on conversations with humans by both understanding and producing speech, hand gestures and facial expressions [33]. The ability of VHs to interact and provide feedback to human users using speech, gestures and facial expressions has made them suitable for various interactive and real-time application [34][35]. Some of these application areas include Engineering, Virtual Conferencing, Training, Games, Education, Virtual Reality, Military and Health Care.

2.2 Application Areas of Virtual Humans

Research has shown that the ability of VHs to provide feedback, generate behavior, portray emotions and make decisions to execute appropriate actions, can have an effective impact on the learning capabilities of the users, thus making these VH interaction interfaces useful across training, pedagogy and education [33] [55] [64] [66] [67] [61] [62] [36] [37]. Earlier research in VH interfaces include a communicative,

Figure 1: Virtual human applications

humanoid agent called Gandalf, developed by MIT in 1998, to guide students in planetary exploration [55]. Users see Gandalf as a hand and face on a small monitor, and interact with Gandalf using scripted speech and gesture. Later in 2000, MIT built REA a virtual real estate agent [33] with the ability to point and describe features of potential properties on sale. "You MD", an interactive museum exhibit, is another example of earlier VH development to improve public health literacy [64]. "You MD" dynamically selects the appearance of the VHs within the interaction based on user information to teach museum visitors about topics in public health, in order to help them make better life choices. The main objective of "You MD" is to increase public health literacy about asthma, melanoma and how to avoid obesity by living a healthy lifestyle [64].

As seen in the above examples, VH development has evolved from simply modeling the physical attributes to creating autonomous agents with the ability to emote emotions, personality, mood, show affect and empathy towards users, and provide cognitive intelligence capabilities like memory, reasoning, decision making and learning. Current development tools like Virtual People Factory (VPF), a web-based tool created by University of Florida, focuses on developing VHs with emphasis on empathy and affect while interacting or communicating with users [66]. The VPF mostly focusses on centralized communication with emphasis on role-playing with an expert in the medical domain to interactively evaluate user's affect, observe their non-verbal behavior, and identify conversational opportunities for empathy [66]. Another such example is the FloRes dialog manager for VHs which uses forward inference techniques, local dialog structure, and plan operators to initiate dialogs between humans and VHs with the aim to support both advanced, flexible, mixed initiative interaction and efficient policy creation by domain experts [67]. Though both VPF and FloRes are good tools to rapidly create VHs and improve dialog between users and VH's, what it lacks is the ability to observe or investigate the effect of the VH's behavior and changing personality or mood on the user behavior communicating within an interactive scenario. Hence both VPF and FLoRes have the potential to be good practice tools for interaction with VH's across various scripted domains with the focus mainly being on communication, rather than tools used to investigate effects of VH's behavior on the interaction and learning outcomes of the user within a specific application domain.

The use of Virtual Humans as Virtual Patients (VP) has increased over the years in health care education too, in response to- 1) the increasing demands on the education of health care professionals and students, and 2) providing students a consistent and controlled environment to practice diagnostic and therapeutic methodologies. In recent years VPs have been implemented to teach bedside competencies of bioethics, basic patient communication and history taking, clinical decision making, and psychotherapy training [36] [37] [38] [61] [62] [63]. VPs allow the learners to take on the role of a health care professionals to develop necessary clinical skills such as making diagnostic and therapeutic decisions [36][37]. These VPs tend to provide a valid, reliable, and applicable representation of live patients and can augment many live actor, simulated, patient programs.

DIANA is one such VP created by Benjamin Lok and his team to help medical students improve their doctor-patient communication skills [38]. VPs can be used to teach caregivers or family members effective interaction techniques to communicate with their aging loved ones or relatives suffering from mental health issues. For example, a VP can simulate an elderly person requiring a great deal of facilitation with the day-to-day activities based on their health needs, like reporting the last time they took their medicine or performed their physical therapy. Researchers at ICT are currently developing VPs with the objective of training social workers in military specific issues and conversations with military personnel on day-to-day life issues after return from service [61][62][63]. VPs can also be used to provide tutorials on clinical skills (neurological exam, breast exam, pelvic exam) and physical examinations, by showing interaction between a physician and patient [37]. The Virtual Patients Tutorial Library at Harvard Medical School is an interactive web-based program enabling users to interact with computer-based patients in simulated clinical encounters [37]. Current research in VP development has evolved to provide clinicians with the fidelity to design their clinical scenarios for practice and learning. On the same lines, Hodges and Dukes at Clemson have proposed to develop a scenario builder tool that allows nurses to build their own scenarios to practice medical diagnosis and communication [60].

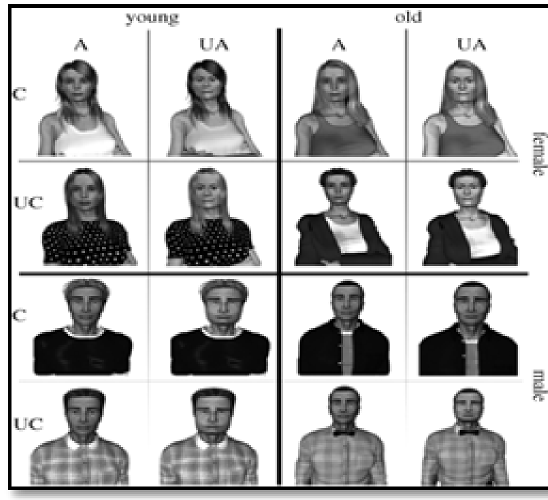
These developments in VH's simulation have evolved to the point where VHs that support rich interactions with people have paved the way towards a new generation of interactive systems for entertainment and experiential learning. Though potential applications of VHs are considerable, there still exists a need to investigate the social impact of VHs on communication and learning outcomes of the users based on the verbal and non-verbal behavior of the VH. The following section investigates and describes the social effect of VHs on user interactions and learning.

2.2.1 Social Effects of Virtual Humans

Pedagogical theories of Vigotsky emphasize the role that social aspects play in successful learning: i.e. knowledge is socially constructed and learning essentially involves sharing and negotiating [40]. Donald Vigotsky also suggests that the rich ability to process various forms of social information, as well as the motivation to do so are essential for human intelligence and existence. Using Vigotsky's theory to develop VHs with the ability to explore and process social information, have been found to be helpful across various social and collaborative platforms, where learning and interaction play a pivotal role to shape user experiences [41] [42] [43].

The social capability of VHs, leads to several benefits for users in pedagogical or training environments. Users can relate to the character as a social and intellectual partner, sharing ideas, questioning and criticizing, like Justine Cassell's Sam who is a

Sam Story-telling Agent - Justine Cassell



Pedagogical Learners - Amy Baylor

Figure 2: VH as social & intellectual partners

story-telling agent designed to tell stories collaboratively with children [44]. Research has shown that users experience affiliation and identify with a pedagogical agent, with these interface agents functioning as social role models [44] [45]. Amy Baylor conducted a study to investigate user's preference in pedagogical learners, especially amongst females to improve female retention in engineering [45]. Studies conducted by Johnson Lewis in investigating interaction tactics with VHs and, studies conducted by Amy Baylor in investigating how VHs help learners with math anxiety, show that learners tend to experience material as less difficult, less intimidating or less anxious in the presence of a pedagogical VH [43] [46]. Users tend to experience less frustration in the presence of an embodied agent [47]. Evidence also suggests an increase in self-efficacy of users interacting with a conversational VH in a task-oriented environment, as investigated by Jeeheon and Baylor to test the usefulness and effectiveness of VHs in a multi-agent collaborative learning environment [48] [49].

To conclude, Nass and Moon have said that people tend to react and attribute human like characteristics of helpfulness, expertise, and friendliness to computers [50]. To affirm on Nass and Moon's theories, researchers have found that users do tend to apply these human attributes of helpfulness, expertise, and friendliness to VHs while interacting and learning with them, as observed in above studies [43] [44] [45] [46] [47] [48] [49]. Though there is always a flip side, where slow reactions and inanimate responses of the VHs can make it frustrating for the people, especially where completing a task is involved.

2.3 Developing Intelligent Agents/Virtual Humans

To simulate intelligent, socially adaptable VHs for a task-based and learning environment various components are involved, starting from the physical embodiment to the cognitive reasoning. The following subsection gives an in-depth review of the various development components involved in VH development, with the main focus on understanding the cognition process.

2.3.1 Components for Developing Intelligent Agents

Building a VH is a multidisciplinary effort, joining traditional artificial intelligence problems with a range of issues from computer graphics to social science. VHs must act and react in their simulated environment, drawing on the disciplines of automated reasoning and planning. To hold a conversation, they must make use of natural language processing research, from speech recognition and natural language understanding to natural language generation and speech synthesis. Additionally, providing human bodies that can be controlled in real time delves into computer graphics and animation. And finally, VH research draws heavily on psychology and communication theory to appropriately convey nonverbal behavior, emotion, personality and actions [34]. To summarize, developing a VH consist of the following closely-linked components:

2.3.1.1 Virtual Human Embodiment

Embodiment involves designing the physical characteristics (skin color, face, hair, etc.), physical behaviors (gestures, physical motion and movements) of VHs that need to be exhibited and developing their communicative functionalities. The VH needs to be responsive; i.e., have the ability to respond to humans and events within the interaction environment. They must be believable; i.e., they must provide a sufficient illusion of life-like behavior that the human user will be drawn to. They must be interpretable; i.e., the user must be able to interpret the VH's response to situations, using the same verbal and non-verbal behaviors that people use to understand one-another [33] [68]. The animation of the body must address a range of technical challenges- like requiring multiple parts of the body to be in motion. For example, a simple wave for a VH requires coordination of the hand, fingers, twisting of the joints in the torso, head/neck and eye movements.

For character animation, a wide range of computer graphic techniques/software and animation packages are available to animate the interactivity of VHs, for instance: walking [69] [70], reaching and object manipulation [71] [72] [73], inverse kinematics [74] [75], and parameterized models [76]. Software packages like Haptek [31], QUIDAM and Ogre3D game engine, use the notion of key frame animations and user-implementable motion generators for character control and animation. Smart-

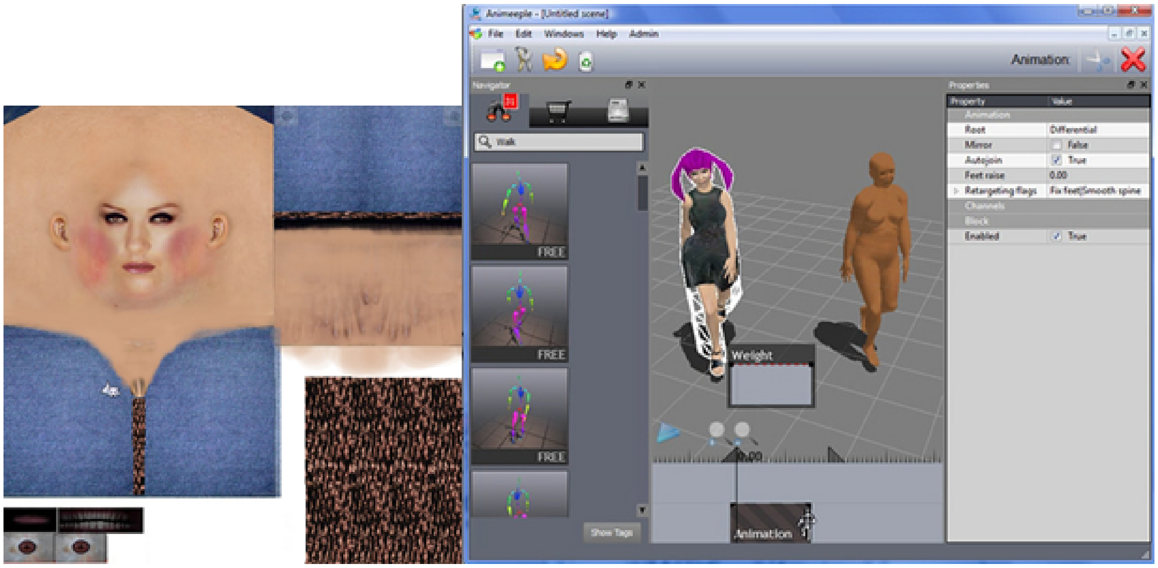


Figure 3: Virtual human embodiment

Body [68], a character animation tool developed by ICT, provides developers with the ability to infuse VHs with locomotion, steering, object manipulation, lip-synching, gazing, and other nonverbal behavior generation capabilities within real time environment.

2.3.1.2 Natural Language Generation

The goal of an interactive environment is to allow VHs to engage in human-like interaction with users and other agents. In order for VHs to successfully simulate real world communication, they need to sound natural and give the impression of being engaged within the interaction. When building a VH dialog module, major consideration is given to the decision of using human recordings vs. synthesized speech. The advantage of using human recordings is performance and the naturalness of the voice quality. But it is expensive to hire a professional actor to record the lines that the VH will utter, and any changes to the script can be an expensive process.

Current state-of-the-art speech synthesizers have reached high levels of naturalness and intelligibility for neutral read aloud speech. However, synthesized speech generated using neutral read aloud data lacks all the attitude, intention, and spontaneity associated with everyday conversations [86]. Comparative studies have been done to study the effects of synthesized speech vs. human speech on the user interactions, as seen in Dickerson's virtual patient dialog model [38] for medical turn taking, and Forbe's intelligent tutoring dialog system [87] for evaluating spoken vs. typed tutoring strategies. Results from Forbe's study show that while in human tutoring, changing the modality from text to speech caused improvements in the learning outcomes and dialog efficiency, in computer tutoring it made less difference. However, in both human and computer tutoring, it was found that changing the modality caused differences in superficial dialog characteristics, and differences in the type of dialog characteristics that correlate with learning [87]. In summary, we can say that there are indeed potential payoffs for adding speech to text-based dialog environments, but more research is needed to fully investigate this potential.

Research has also been done to compare professional vs. amateur human recordings and limited-domain vs. general-purpose synthesized voices, like in the virtual human dialog system called SimCoach [88], which aims to motivate military personnel and family members to seek information and advice with regards to depression and post-traumatic stress disorder. Results from SimCoach suggest that a professional human voice can supersede both an amateur human voice and synthesized voices. Also, a high-quality, general-purpose voice or a good limited-domain voice can perform better than amateur human recordings. No significant differences were observed between the performance of a high-quality general-purpose voice and a limited-domain voice, both trained with speech recorded by actors [88].

As seen, research in the area of natural language generation is still limited and computationally challenging. There is still a need for further exploration.

2.3.1.3 Cognitive Artificial Intelligence

The term Cognitive Artificial Intelligence (AI) has been coined with the goal of implementing aspects of human intelligence and behavioral models in computers [83]. Cognitive AI, also termed as Embodied AI, defines the integration of various cognitive theories such as, symbolic cognition (language, planning, high-level deliberation) with sub-symbolic processing (perception, analogical reasoning, neural learning and classification, memory retrieval etc.), and action regulation to form a broad intelligent cognitive architecture [83]. Thus, any cognitively intelligent architecture should consist of the following attributes:

- Set of suitable events.
- Evaluation method to establish goals and identify adverse events.
- Computation model to represent environmental situations and events.
- Protocol memory model to store past situations and events.
- Reinforcement learning mechanism.
- Recall model for memory to recollect memory contents based on current environmental situation and needs.

- Decision-making component to deduce an appropriate action to control current events or the task that needs to be achieved.

- Computational model to generate set of appropriate actions to realize a particular task.

- Planning, classification and problem solving model, to actively construct ways to reach a goal situation.

- Mechanisms for reflection, reorganization and abstraction of existing memory content.

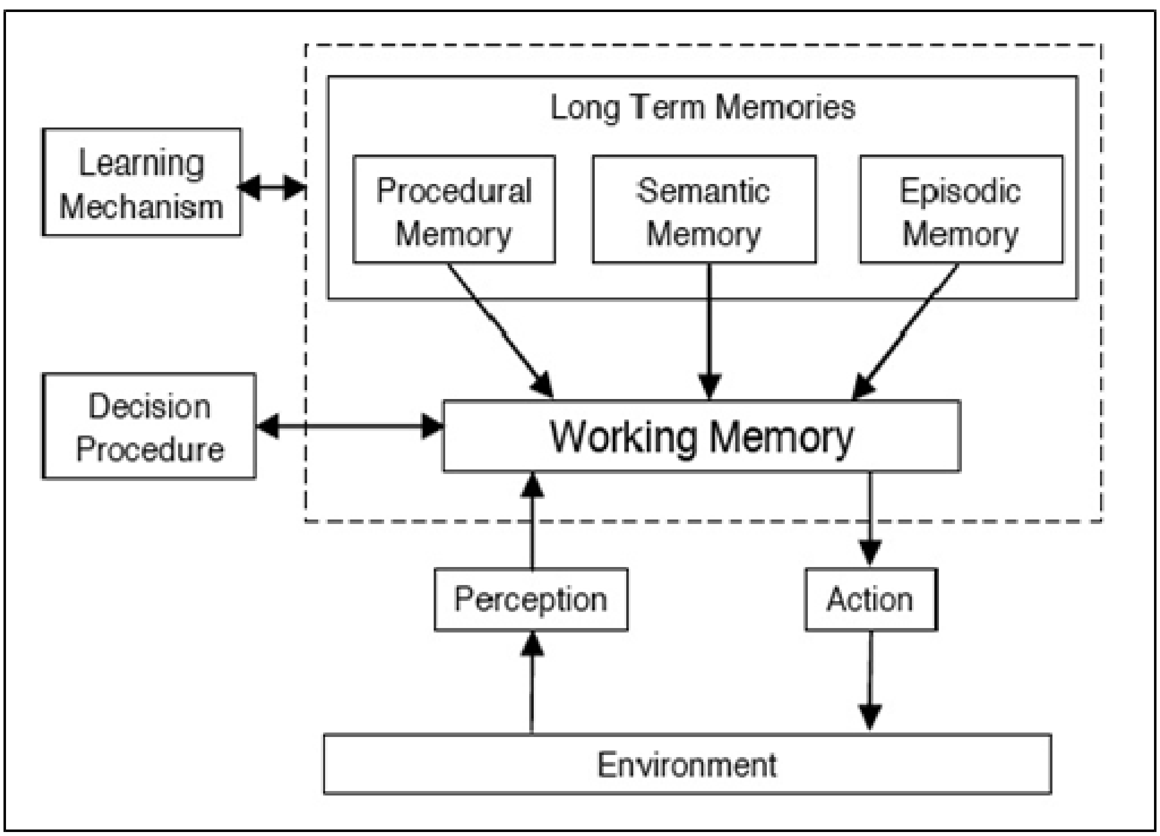


Figure 4: SOAR architecture

Earlier Cognitive Architectures: Various cognitive architectures have been defined over the years to drive the behavior generation of the VH, based on the mental state and world knowledge of the agent. Earlier work in cognitive architecture includes the SOAR architecture (see Figure 4), which is based on symbolic reasoning and deduction, and is suitable for dynamic environment where tasks and communication between the user and the VH is scripted by experts, rather than real time applications where the tasks/goals can change based on the user requirement [53]. Another such earlier architecture is the reactive architecture proposed by Brooks [2], where behaviors are represented as simple if-then rule structures computed in a hierarchical manner where each agent has precedence over another. Finite State Machines (FSMs) are commonly used reactive architectures where behaviors are arranged as hierarchical tasks (see Figure 5) and a new state results from the measured effects of the current state condition [3]. More recent approaches have moved towards more

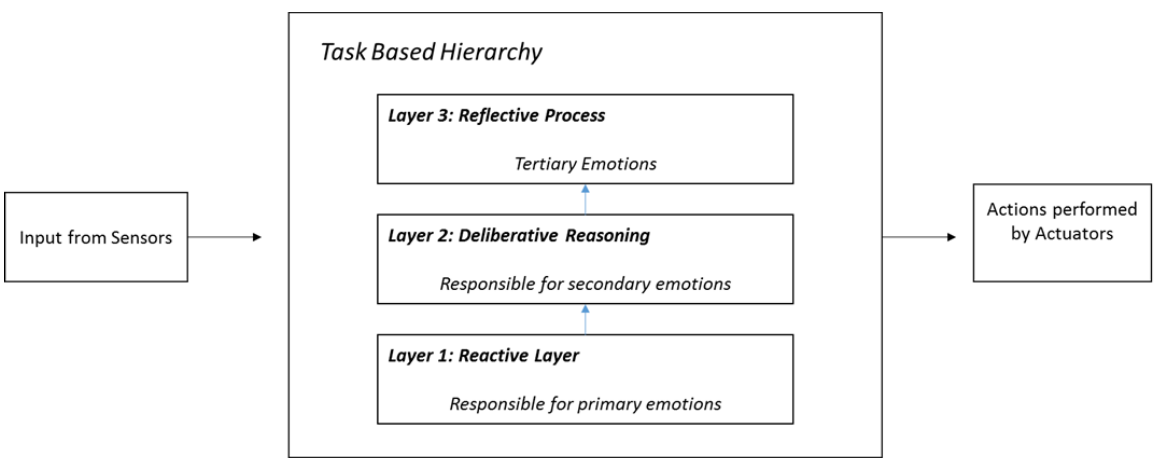


Figure 5: FSM: Reactive architecture

abstract reasoning frameworks, largely building on traditional artificial intelligence techniques. In FLAME (see Figure 6) an abstract agent architecture by El Nasr and colleagues [79], Markov's Decision Process (MDP) is used to provide a general framework for characterizing the desirability of actions and events. The FLAME architecture determines the behavior of the VH, based on an underlying appraisal theory which is responsible for making decisions and generating an emotional state. Current

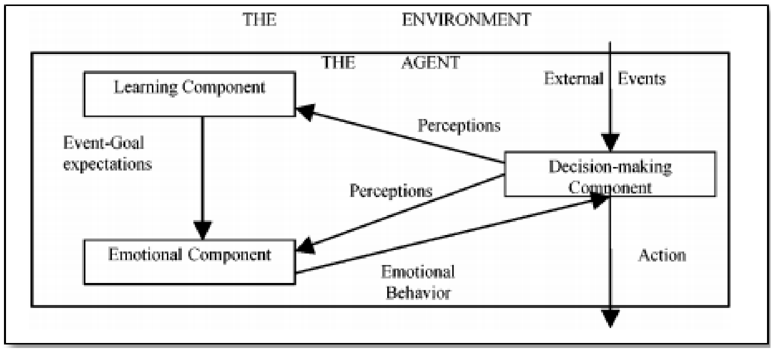


Figure 6: FLAME: Abstract agent architecture

Cognitive Architecture: Current cognitive architectures are based on two main computational models (see Figure 7): the appraisal model (responsible for calculating the desirability and expectedness of an event) and the emotion model (responsible for mapping the appraisal variables onto an emotion space).

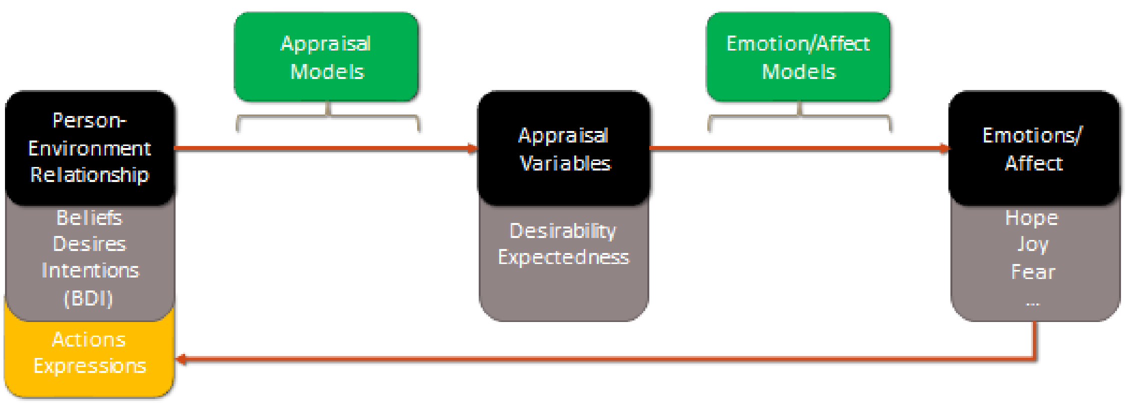


Figure 7: Structure of current cognitive architectures

- Appraisal Model: Appraisal is the process by which a person assesses their overall relationship within the environment, based on their current conditions, past events leading up to the current state, and future prospects for achieving the goal [84] [85]. Usually in an appraisal process, events are appraised based on the symbolic representation of agent's belief (B), desire (D) and intention (I), where: *Beliefs* are facts representing what an agent believes about the world, i.e. an agent's representation of the state of the world, *Desires* are goals or some desired end states (an agent can have multiple desires) and *Intentions* refer both to an agent's commitments to its desires (goals) and its commitment to the plans selected to achieve those goals (an agent's intention need to be consistent) [4]. One such architecture based on the underlying BDI model is the Elliott's [78]

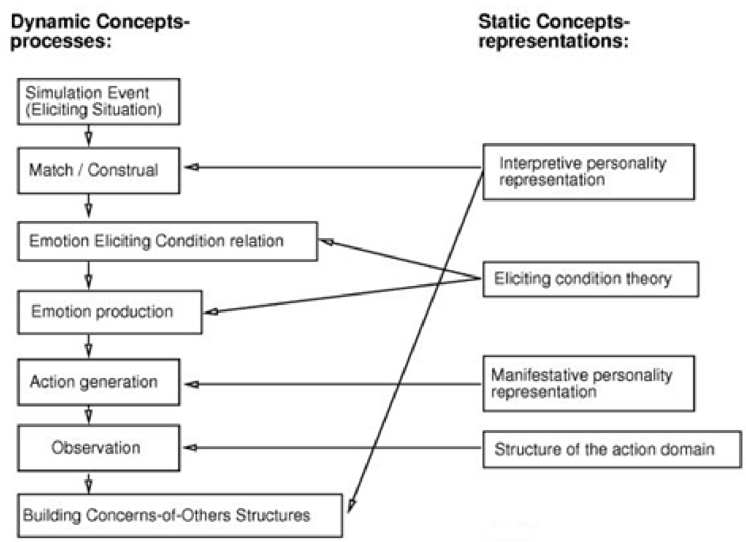


Figure 8: Elliot's affective reasoner

Affective Reasoner (AR) (see Figure 8), which uses domain specific rules to appraise events. For example, in a football match scoring a goal is desirable if the agent favors the team that scored, otherwise the agent would perceive the event of scoring a goal as undesirable. Additionally in AR, agents use a case-based heuristic classification system to reason about the emotions of other agents and to build representations of those other agents' personalities that will help them predict and explain future emotional episodes involving those observed agents.

Appraisal is also said to trigger cognitive responses, often referred to as coping strategies- e.g., planning, procrastination or resignation- feeding back into a continuous cycle of appraisal and re-appraisal [84] [85]. People respond to events differently depending upon their appraisal of the current event [77]. Coping is a process that determines how one responds to the appraised events. For example, events appraised as undesirable but controllable motivate people to develop and execute plans to reverse these circumstances. On the other hand, events appraised as uncontrollable lead people towards denial or resignation. The EMA architecture (see Figure 9) by Gratch and Marsella, is a decision based theoretical planning model that allows agents to appraise events and re-evaluate their current actions by applying a coping strategy based on probabilistic reasoning [65]. Thus any cognitive architecture that claims to support the full range of appraisal and coping strategies must minimally satisfy and map onto the following appraisal variables:

– To capture the constructive and interpretative nature of the processes and represent intermediate knowledge states.

– To reason about relevance and desirability, the architecture must represent preferences over outcomes (Relevance, Desirability).

– To make casual attributions, the architecture must represent some notion

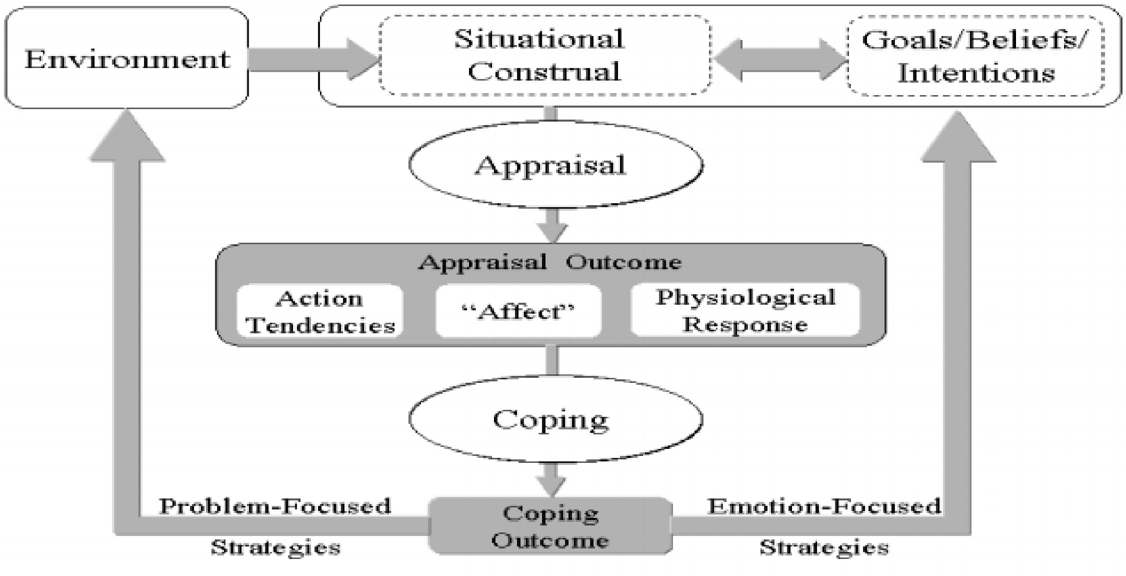


Figure 9: EMA agent architecture

of causality and agency (Casual Attribution).

- To reason about likelihood, unexpectedness and changeability, the architecture must represent causal factors influencing events, future possible outcomes and interactions between possible outcomes (Likelihood, Unexpectedness, Changeability).

- To reason about the urgency, the architecture must represent temporal constraints, event duration, and, partial goal achievement (Urgency).

- To reason about controllability, the architecture must represent the extent to which events can be controlled (Controllability).

- To reason about social power, the architecture must have some representation of coercive relationships between agents (Power).

- To reason about adaptability and to support emotion focused coping strategies, the architecture must be open to subjective reinterpretation (Adaptability).

- To reason about ego-involvement, the architecture must support some notion of how central a desire is to the agents self-concept (Ego).

- Emotion Model: Also known as the Affect model is responsible for mapping the appraisal variables into some behavioral or cognitive change (emotional state). Emotions can be directed outward into the environment or inward, shaping a person's thought process. Reflecting on this, cognitive architectures determine how affect alters the nature or content of these cognitive processes. Affect is a representation of the agent's current emotional state. This could be a discrete emotion label, a set of discrete emotions, core affect (i.e., a continuous dimensional space), or even some combination of these factors. An important consideration in representing affect, particularly for systems that model the consequences of emotions, is for this data structure to preserves the link between appraisal factors and emotional state [84][85].

  To generate the cognitive behavior (affect or emotional state) of an agent from a set of appraisal variables (as defined by the appraisal process above), most of the modern cognitive architectures make use of one of the following psychological theory of emotion generation:

  1. Lazarus Theory: It defines emotions according to 'core relational themes' which are intuitive summaries of the 'moral appraisals' (e.g. of relevance, goal conduciveness) involved in different emotions. These themes help

define both the function and eliciting conditions of the emotion. They include: Anger (a demeaning offense against me and mine), Fear (facing an immediate, concrete, and overwhelming physical danger), Sadness (having experienced an irrevocable loss), Disgust (taking in or being too close to an indigestible object or idea) and Happiness (making reasonable progress toward the realization of a goal) [90].

2. Scherer Theory: It defines appraisals of events based on an invariant sequence [91]. This sequence model assumes the appraisal process to be constantly operative, with evaluations constantly performed to update the agent's information about an event or situation. The Scherer theory results in an unfolding sequence of emotional responses that includes: Joy, Fear, Anger, Sadness, Disgust, Shame and Guilt.

3. Frijda Theory: The center of Frijda's theory is the term concern [92]. A concern is the disposition of a system to prefer certain states of the environment and of the own organism over the absence of such conditions. Concerns produce goals and preferences for a system. If the system has problems to realize these concerns, emotions develop. As substantial action tendencies, Frijda defines the following associated emotions in parentheses: Approach (Desire), Avoidance (Fear), Being-with (Enjoyment, Confidence), Attending (Interest), Rejecting (Disgust), Non-attending (Indifference), Anger (Attack/Threat), Interrupting (Shock, Surprise), Dominating (Arrogance) and Submitting (Humility, Resignation).

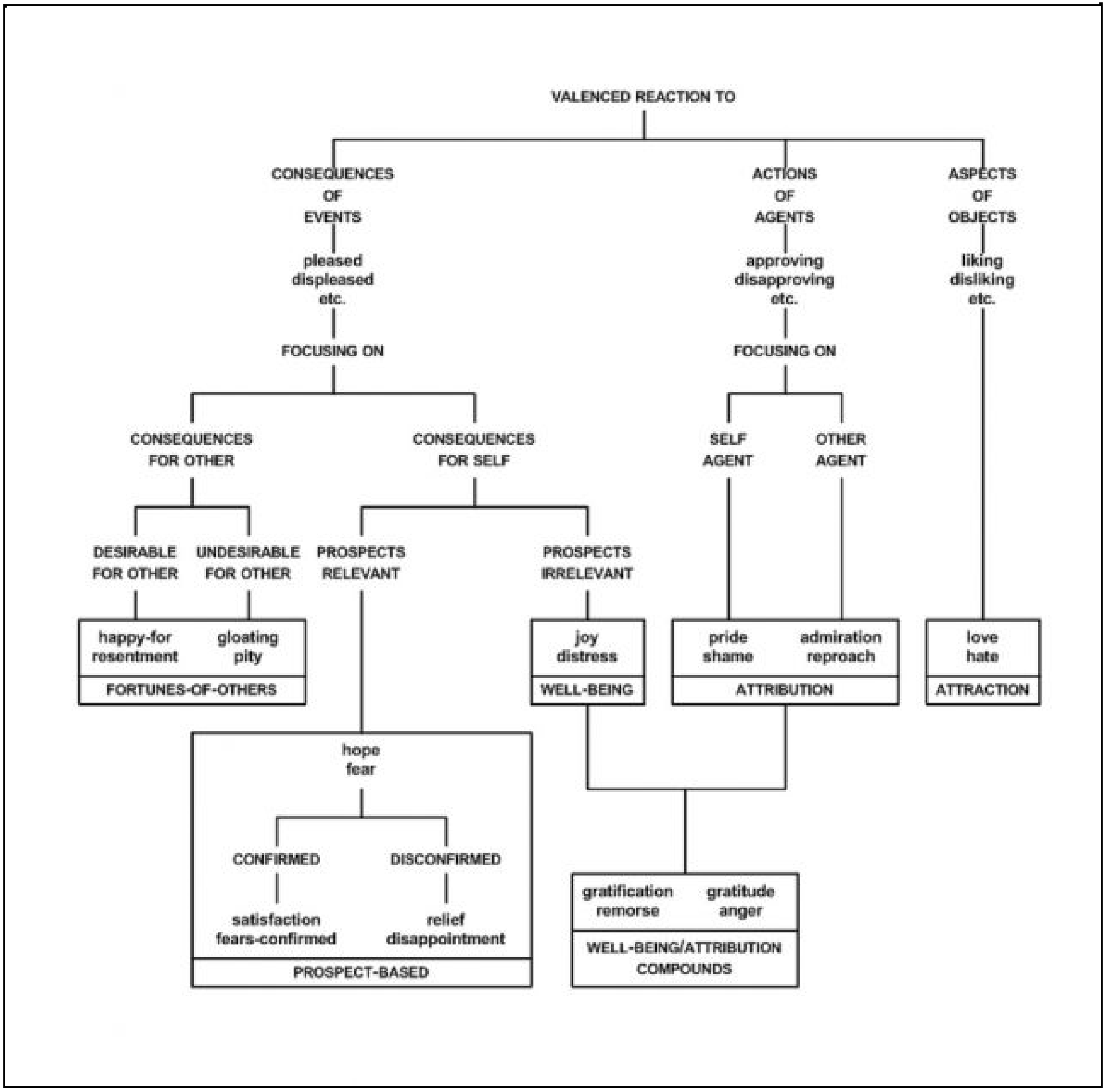


Figure 10: OCC emotion model [30]

4. OCC Theory: The OCC model of emotions by Ortony, Clore, and Collins [30] defines emotions as a valenced reaction to events, agents, and objects, and considers valenced reactions as a means to differentiate between emotions and non-emotions. According to the OCC model, all emotions can be divided into three classes, six groups, and 22 types as seen in Figure 10. The OCC model constitutes of a goal, standard, and attitude-oriented

emotion appraisal structure. Thus making it easier for applying natural language processing (NLP) tools for the identification of emotion-inducing situations (e.g., event/action), the cognitive state of the user (usually expressed by adjectives and adverbs), and the variables causing emotion (e.g., real-world knowledge about something or somebody etc.).

To summarize many modern cognitive architectures (see Figure 11)use the underlying computational theory of appraisal and coping to appraise an agent's situation and event to generate a set of appraisal variables which are then mapped onto an appropriate cognitive state using an emotional/affect model, which manifests into an appropriate behavioral state displayed in the form facial expressions and gestures performed by the agent.

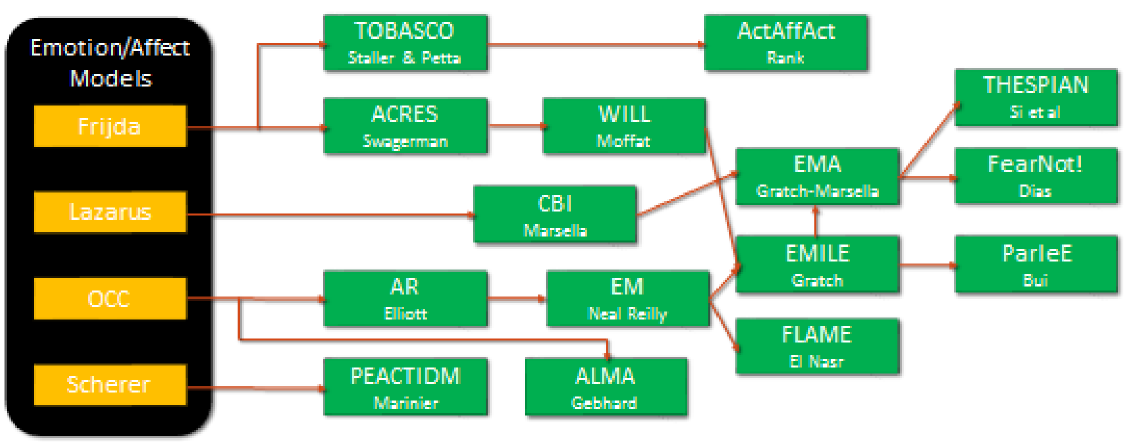


Figure 11: Modern cognitive architectures

2.4 Limitations with Existing Architectures

To create a truly intelligent agent, a cognitive architecture should consist of an Appraisal Model mapping onto an Emotion Model. But most of these cognitive architectures do not address this direct mapping, as seen in the Affective Reasoner[78] and the ALMA[93] model where mapping between appraisal variables and emotions are scripted based on the task environment (domain-specific appraisal rules). Psychological theories suggest that agent behavior results from a span of motivationally related concepts ranging from the basic needs and drives to highly abstract concepts like cultural background, family, education and so on. Most of the modern cognitive architectures fail to take these concepts into consideration while driving an agent's behavior. For example, the EMA [65] agent architecture though allows for coping strategy, but it's emphasis is more towards goals and goal processing.

To conclude all the current agent architectures have the ability to generate behavior based on perception, decision-making and action control but fail to encode agent's individualistic characteristics like cultural beliefs, personal history, and past experiences that shape their behavioral and emotional outcomes.

CHAPTER 3: PROBLEM DEFINITION & CONTRIBUTIONS

3.1 Problem Summary

Prior cognitive architectures like AR (Elliot's Affective Reasoner) and ALMA help determine an agent's mental state by appraising events based on desirability, but these appraisals use domain-specific rules which are scripted to match the task environment in which the agent is interacting. Even agent architecture like EMA, with it's ability to re-appraise and re-evaluate the events using some kind of coping mechanism, is mainly focused towards end goal and goal processing states while computing the agent behavior. The most common theme across all the architectures seen in Figure 11 is generating behavior and cognitive state based on the agent's perceptions about its environment (task related), and its decision-making based on short term goals and end goal. Though these can be helpful in task-based interaction, but has found to be less effective where learning in involved, compared to other learning-focused architectures. As psychological theories suggest that human behavior is a manifestation of various motivational concepts like an individual's need and drives and abstract concepts like cultural beliefs, personality, education and prior experiences, there is a need to explore the feasibility of incorporating these abstract motivational concepts within an architecture to generate a truly intelligent agent.In artificial intelligence, an intelligent agent is an autonomous entity which has the ability to observe through sensors and act upon its environment using actuators (perform actions) and is able to direct its activity towards achieving goals (provide some kind of 'rationale' based on decision making component). Hence my first research goa is to:

1. To create a cognitive architecture that has the ability to appraise (make decisions about) an event based on agent's personal history that incorporates agent's memory/knowledge, cultural belief's, agent's personality and agent's mood.

To this effect, an agent architecture was developed called CMAA where agent behavior is driven by- 1) personality of the agent (across Big Five Personality dimensions [8]- Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness (see Section 5.2.2.2)), 2) Mood of the agent across Mehrabian's mood dimensions [89]- Pleasure, Arousal, Dominance (see Section 5.2.2.2)) and 3) Cultural Belief of the agent (across Hofstede's cultural dimensions [29]- Power Distance Index, Individualism, Uncertainty Avoidance Index, and Long Term Orientation (see Section 5.2.2.2)). These three attributes factor into the appraisal model, to compute a set of appraisal variables. These appraisal variables are then mapped into the OCC emotion model, resulting into 22 diferent emotions.

One of the important functionality associated with VHs is their ability impact learning outcomes within a social task-based environment. Various VH applications like Virtual People Factory, FloRes and SimCoach have shown to be more focused towards creating streamline communication and using communication cues to determine the user's emotional and mental state within an interactive training environment. This gives rise to my next problem definition:

2. To investigate the effects of developing an autonomous intelligent VH on user's engagement, satisfaction, success, and learning in an interactive training environment.

To evaluate these factors an interaction interface with a virtual patient portraying symptoms of mild TBI (Traumatic Brain Injury) was developed, as a diagnostic training tool to standardize the TBI evaluation across VA facilities.

3.2 Research Contributions

This section highlights my research contributions towards the completion of my dissertation project as follows:

1. Developed an Agent Architecture called Culturally Modified Agent Architecture (CMAA) that allows the generation of autonomous intelligent VHs based on three main factors: mood, personality and cultural beliefs of the agent (captures based on agent's past experiences and personal values).

2. The feasibility of the CMAA architecture was tested by developing and implementing a VH in a clinical setting as a Virtual Standardized Patient (VSP), portraying an OEF (Operation Enduring Freedom)/OIF (Operation Iraqi Freedom)/OND (Operation New Dawn) Veteran exhibiting symptoms of mild TBI (Traumatic Brain Injury) as a training tool to: 1) Test the validity and believability of the VH prototype as veteran portraying symptoms of mild TBI and 2) test the effectiveness of using the VH as a training tool to practice diagnostic evaluation and improve communication between a patient and a provider in a clinical setting.

CHAPTER 4: PRELIMINARY WORK: DR. CHESTR SHOW

4.1 Introduction

VHs have proven effective in several learning applications that include: presenting information, providing education and engaging in social conversations. VHs as learning pedagogies have shown to contribute significantly towards increasing user engagement and satisfaction within pedagogical applications. This is due to their ability to portray believable and realistic human characteristics, i.e. portray personality and emotions. Thus my research focuses on developing VHs with personality and emotions as they tend to influence the behavior of human characteristics within a learning environment. The initial research objective was to develop VHs with the ability to interact naturally and spontaneously using speech, personality and emotions. As a result Dr. Chestr a virtual game show host was designed and implemented, infused with a unique personality to promote user engagement and enjoyment. Dr. Chestr was designed to test users with questions about the C++ programming language, its interface allows user to communicate with Dr. Chestr using the most natural form of interaction- speech. This chapter describes the design and implementation of Dr. Chestr and its underlying agent architecture for developing a personable VH. The chapter concludes with results from the evaluation study conducted to 1) evaluate Dr. Chestr's personality and to 2) test the feasibility of using the Dr. Chestr interface as a learning tool within an educational setting.

4.2 Dr. Chestr Game Show

Dr. Chestr (see Figure 12) is a computerized virtual game show host that was developed and deployed as a study tool for students learning basic programming concepts in the introductory courses of Computer Science (ITCS 1212). Dr. Chestr was designed to quiz students on the C++ programming concepts covered during their lecture and lab sessions. The questions in the study tool are posed in a multiple choice or (T)rue/ (F)alse format, and students get to interact (i.e. answer questions) with the Dr. Chestr interface using two communication modalities speech and mouse clicks.

Figure 12: Game show host Dr. Chestr

4.2.1 Overview of the Game Show

The game show host, Dr. Chestr initiates the game by introducing himself and his functionality. Next he guides the user through the logistics and rules of the game. Each game segment consists of 10 questions that include both multiple choice and true/false. Dr. Chestr recites the questions and possible answers to the users. To improve understandability, the questions and answers are also displayed on the screen. Once Dr. Chestr poses the question, there is no time limit on the user's response. Only verbal responses are accepted, initiated by the depression of the 'push to talk' button. Users can answer question using the NATO phonetic alphabets 'alpha','bravo','charlie' and 'delta' to improve speech recognition. Dr. Chestr responds in affirmation or negation after the user replies to the question. There is a fifty-fifty option available to the users providing assistance, thus increasing user's chances of selecting the correct answer from 25% to 50%, by reducing the number of choices from four to two. The user can choose the 'fifty-fifty' option only once during every game segment. At the end of 10 questions, the user's performance is displayed. Dr. Chestr's affirmation/negation of user's choice is one of the main avenues through which his personality is displayed.

4.2.2 System Architecture

Dr. Chestr's appearance is implemented using Haptek's People Putty[31] and his gestures are designed using Haptek's Figuremaker[31]. The Dr. Chestr game is a web-based study tool that runs on Internet Explorer via a Haptek player. The study tool uses Microsoft's TTS (Text-to-Speech) Engine to synthesize user's spoken dialog and AT&T natural voice to generate scripted speech for Dr. Chestr to communicate with the users. Dr. Chestr's speech and gestures are generated during real-time via Haptek's plug-in using Javascript. Dr. Chestr's knowledge repository consists of questions and answers categorized into eight different chapters based on the data source (online resource and textbook) provided to us by the instructors. Users' responses to the questions being asked are matched to the stored answers using an information retrieval algorithm for answer matching.

4.2.3 Agent Architecture

To drive the behavior of Dr. Chestr (verbal output and non-verbal output) an agent architecture was developed (see Figure 13) consisting of three main components:

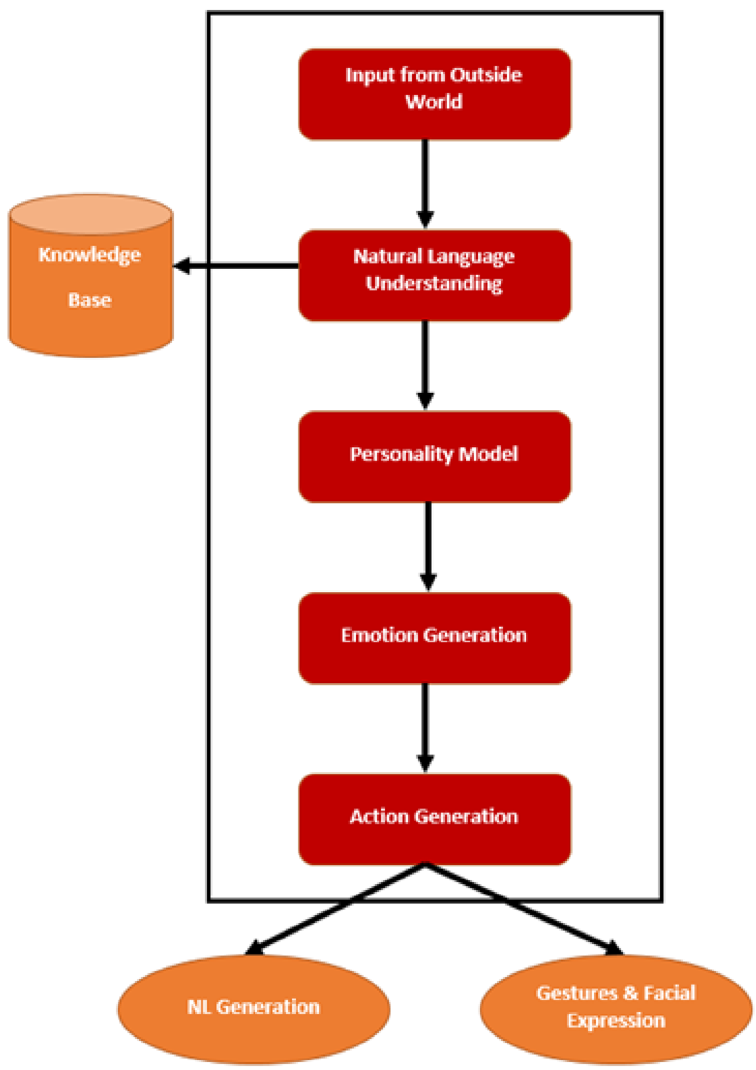


Figure 13: Dr. Chestr agent architecture

1. Natural Language Understanding: Responsible for generating a question and synthesizing users response based on the domain knowledge represented in our knowledge base.

2. Personality Model: Responsible for applying a personality to Dr. Chestr. Personality is defined as a set of characteristics that a person possesses making him or her unique in terms of cognition, behavior and motivation in various situations [8]. To model Dr. Chestrs personality I have utilized The Five Factor Model (FFM) by Robert McCrae and Oliver John [9]. The FFM was chosen due to its ability to breakdown the personality traits of an individual in terms of five basic dimensions (see Table 1): Extraversion (extrovert vs. introvert), Agreeableness (cooperative vs. antagonistic), Conscientiousness (self-discipline), Neuroticism (anger, depression) and Openness (learn, appreciate and experience new things) [8] [9]. Dr. Chestrs personality is represented using the following equation:

$$E(0.9) + A(0.3) + C(0.2) + O(0.4) + N(0.0) \tag{1}$$

Weights were assigned to each dimension of personality on a scale of 0-1 depending upon their significance in defining Dr. Chestrs personality. Dr. Chestr was designed to represent a young, dynamic professional male with a warm, outgoing and engaging persona. The weightings that define an engaging and outgoing persona were chosen based on the adjectives defined in Table 1:

   - Talkative, skilled in humor, aggressive, facially and gesturally expressive but not gregarious - 0.9

- Skeptical, critical, not overly trusting and agreeable - 0.3
- Self-indulgent and slightly unethical in responses - 0.2
- Responses unconventional, original - 0.4
- No depression and anger - 0.0

A sliding scale was used to map the adjectives and definers to each factor within equation 1. The resulting weightage from equation 1 was then supplied to the action generation module.

Table 1: Five Factor Model (FFM) of personality

Five Factor Model	**Adjectives**	**Factor Definers**
Extraversion (E)	Active	Talkative
	Assertive	Skilled in humor
	Enthusiastic	Facially, gesturally expressive
	Outgoing	Behaves Assertively
	Talkative	Gregarious
Agreeableness (A)	Trusting	Trustful
	Appreciative	Not critical, skeptical
Conscientiousness (C)	Planful	Not self-indulgent
	Responsible	Behaves ethically
Openness (O)	Original	Judges in unconventional terms
Neuroticism (N)	Anxious	Thin-skinned

3. Action Generation Module: Responsible for generating a set of verbal (speech) and non-verbal output (emotions and gestures). The resulting weightage from the personality model was factored into a set of if-then rules to generate an appropriate emotion (pleasure, surprise, sadness and anticipation), set of gestures (winking, pointing finger, clapping, smiling and a sad nod) and verbal feedback from the knowledge repository to the user.

Hence the development of an extroverted, personable, intelligent, slightly conceited and sarcastic virtual human in Dr. Chestr was accomplished (see Figure 14) through

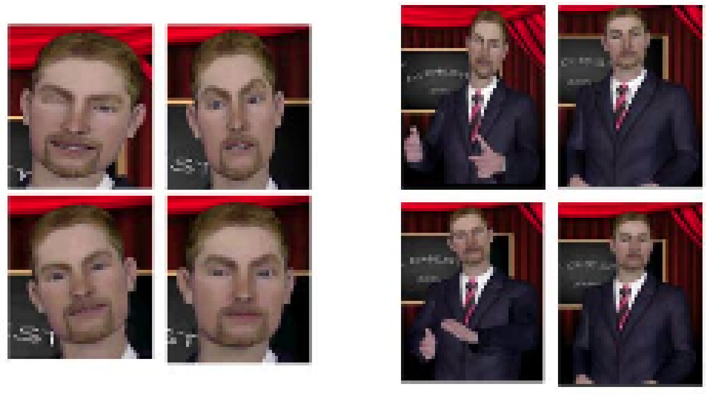


Figure 14: Dr. Chestr's personality

the combination of the afore-mentioned components of speech, personality and action generation.

4.3 Evaluation Study

A usability study was conducted, to test the effectiveness of Dr. Chestr as a study tool. 250 students enrolled in the introductory computing course of ITCS 1212 were recruited as subjects for the evaluation study. A controlled experiment was conducted in which each subject interacted with the Dr. Chestr interface for one game segment (10 questions). To reduce the causal effects of other factors, the following controls were applied:

- All participants sat in the same chair and used the same PC with the same equipment (microphone, speakers) for the study.
- All participants completed the same task of answering 10 questions. The questions were randomly selected from the knowledge base.
- Identical procedure's were followed for each participant. What followed was

the presentation of the consent form, the pre-experiment survey and lastly the PANAS (the Positive and Negative Affect Score) was administered. Each subject was then given a login and password and instructed to initiate interactions with Dr. Chestr.

- Upon completion of the game segment, another PANAS test was administered followed by the post-experiment questionnaire.

- All participants were instructed to not discuss the experiment with their classmates to ensure that all participants had equal knowledge of the study.

4.3.1 Results

A pre-experiment questionnaire (see Appendix A) was given to the subjects at the start of the study session. The pre-experiment questionnaire collected demographic information about the subjects including their familiarity with computers. At the end of the experiment a post-experiment questionnaire (see Appendix A) was given to the subjects to evaluate their interaction with Dr. Chestr. The study also measured user's perception of Dr. Chestr's human-like characteristics, including appearance, behavior, voice, personality, facial expressions and gestures. The evaluation study was conducted to evaluate the following key questions:

- Whether users were able to identify Dr. Chestr's personality?

- Did Dr. Chestr make the interaction more enjoyable?

- Did the users like Dr. Chestr?

- Would the users interact with Dr. Chestr again?

- Were there any changes in users mood before and after interacting with Dr. Chestr?

4.3.1.1 System Usability Scale (SUS)

Usability of the system was measured using the System Usability Scale developed by the Digital Equipment Corporation[94]. Overall the system exhibited a high degree of reported usability (M = 78.00813 SD =13.05317) with a reported highest usability score of 100.0 and lowest score of 30.0.

4.3.1.2 Test of Positive and Negative Affect

Participants positive and negative affect were measured prior to the experiment session and also immediately after the experiment session using the Watson, Clark and Tellegan Positive and Negative Affect Test termed as PANAS [32]. The test consisted of 10 positive affect questions and 10 negative affect questions, measured on a Likert Scale (1 = no affect seen or slight affect, 5 = extreme affect). PANAS has it‘s own method of calculating the mean value, across an average value of 40. Analysis

Table 2: Descriptive statistics for positive and negative affect score for pre- and post-experiment condition

Condition	**Mean**	**Std. Deviation**	**N**
PANAS Positive Pre-Test	32.44	6.75	133
PANAS Positive Post-Test	33.96	8.48	133
PANAS Negative Pre-Test	13.57	3.63	133
PANAS Negative Post-Test	12.19	3.32	133

of the PANAS positive score revealed that overall positive affect increased from the pre-experiment sessions (M = 32.44, SD = 6.75) to the post-experiment sessions (M = 33.96, SD = 8.48), though the effect size was small. Analysis of the PANAS negative score revealed that there was a notable decrease from the pre-experiment sessions (M = 13.57, SD = 3.63) to the post-experiment sessions (M = 12.19, SD = 3.32) (see Table 2). A t-test was performed to compare the mean scores for Positve and Negative Affect across the pre-experiment session and post-experiment session. Though notable difference was observed across the Negative Affect, the t-test results didn't indicate any significant difference ($p > 0.05$).

4.3.1.3 Evaluation of Dr. Chestr

A qualitative evaluation of Dr. Chestr was done post-experiment, where users evaluated Dr. Chestr on a Rating Scale of 10 - 20 (20 - Strongly Agree and 10 - Strongly Disagree). Overall, the evaluation scores were quite high. Majority, 78%, of the users found Dr. Chestr's appearance realistic and his personality identifiable. Users thought that Dr. Chestr was outgoing, approachable and engaging (M = 19.92, SD = 4.61). 22% of the users found Dr. Chestr to be funny, witty and entertaining. Users indicated that they found interacting with Dr. Chestr intuitive, and his response fast and clear (M = 17.53, SD = 2.46). Some of the comments made by users included:

"More realistic than anticipated."

"He really did have the look of a game show host."

"He was surprising at times with his funny random comments."

"Real game show experience, very detailed even down to facial hair."

Users also indicated that Dr. Chestr would be a useful study tool for Introductory Courses in Computer Science and they would like to use Dr. Chestr again (M = 8.22, SD = 1.71). When asked whether they would have performed worse or same with another method of practice? 43% users felt they would have done the same, while 57% of the users indicated that they would have done worse. Finally, 80% users enjoyed interacting with Dr. Chestr and would use it again, if made available.

4.4 Conclusion

Results from the evaluation study showed us that users were able to identify with Dr. Chestr's personality. Dr. Chestr was able to enhance user experience and made the interaction more enjoyable. Though not statistically significant, the increase in users PANAS scores after interacting with Dr. Chestr can have significant implications in the recruitment and retention of computer science students. The successful development of Dr. Chestr with an identifiable personality provided the foundation for the development of VHs that can portray personality and emotions to enhance user experience and learning outcomes.

4.5 Problem Definition & Research Objective

The preliminary work showed that it is possible to create autonomous VHs using an intelligent agent architecture with the ability to portray personality, mood and emotions that are adaptable across different domains and application areas. Psychological research suggests that human emotions are influenced by one's culture, attitude, ethics, personality and values [12] [13]. An intelligent agent architecture should be able to incorporate these cognitive theories to better replicate human behavior. As a result, an intelligent agent architecture was proposed with the ability to encode one's personality, his/her cultural beliefs and mood to determine an emotional state to drive the behavior of the agent.

CHAPTER 5: CULTURALLY-MODIFIED AGENT ARCHITECTURE (CMAA)

5.1 Introduction

The design and implementation of believable intelligent agents, has been typically addressed from two interrelated perspectives within cognitive science. On one hand, psychological models of human cognition try to explain how human behavior is produced. On the other hand, computational models implemented in artificial agents try to replicate to some extent human-like behavior. My research goal is to create a general computational model underlying human behavior and emotions that supports the development of automated intelligent agents with cognitive abilities to improve learning outcomes in task-based environment. As seen in the the Literature Survey, to develop an agent architecture, cognitive AI supports two underlying computational models: the appraisal model (responsible for appraising events and situations from agent's perspective) and the emotion model (responsible for generating a set of emotions to drive the decision-making component and behavior of the agent). Most of the agent architectures discussed in the preceding sections fail to address this direct mapping between the appraisal model and the emotion model (i.e. the mapping is hard-coded or scripted using set of rules). To develop a truly autonomous intelligent agent architecture it is necessary to address this issue, by incorporating an automatic mapping between appraisal model and the emotion model.

Also emotions play a powerful, central role in our lives. They impact our beliefs, inform our decision-making and in large measures affect our behavior to our surrounding world. Human emotions are directly influenced by one's culture, attitude, ethics, personality and values. Thus the Culturally Modified Agent Architecture (CMAA) was developed with the ability to appraise events based on one's cultural beliefs, personality, mood and map them onto an emotional state using automated mapping rules. This chapter discusses the design and implementation of the CMAA agent architecture.

5.2 Culturally Modified Agent Architecture (CMAA)

An agent architecture called, Culturally Modified Agent Architecture (CMAA) was developed to afford the generation of autonomous intelligent Virtual Humans (see Figure 15), based on 3 categories 1) agent's belief (past history and personal experiences), 2) agent's personality and 3) mood of the agent. To give a brief overview of the CMAA agent architecture in Figure 15 - I start with the Knowledge Base, where agent's knowledge about the world is stored. Then we have a Sensor which captures the input from the agent's working environment (here working environment refers to the application where the agent resides). This input from the Sensors is then passed onto the Symbol Translation, which is responsible for encoding and computationally representing the input (termed as meta-data). This meta-data is then passed onto the Appraisal Process, where the Motivational System is invoked first. This Motivational System is responsible for calculating the desirability of the event that is taking place (i.e. how the event is perceived from the agent's perspective).

Once the desirability is calculated, the Motivational System then captures the value associated with personality, mood and cultural dimension. The P (Personality) - M (Mood) - C (Cultural Dimension) value set along with the desirability value are then mapped onto a set of appraisal variables defined within our Appraisal Process. Once the appraisal variables are captured, the Deliberation Process is invoked which is responsible for invoking the underlying emotion model (I am using OCC emotion model). In the Emotion Model, the appraisal variables are mapped onto a set of emotions based on the rules defined by the OCC emotion model. This emotion set along with the emotion intensity are then passed to the intention structure, where the appropriate action is chosen. By action we refer to the set of animations (includes both gesture and facial expressions) are chosen. This information is stored within the memory associated with the agent, i.e. update on the Knowledge base is performed. Finally, the Effectors are activated which are responsible for the physical simulation of the resulting action performed by the agent.

The two main factors of CMAA, which govern an agent's behavior, are: the appraisal process (responsible for appraising events as desirable or undesirable based on mood, personality and agents belief) and the deliberation process (responsible for generating an emotional state and actions) [25]. The following subsections explain the working and in-depth implementation of the Appraisal the Deliberation Process of the CMAA agent architecture. The design and implementation of different components of CMAA are explained as follows:

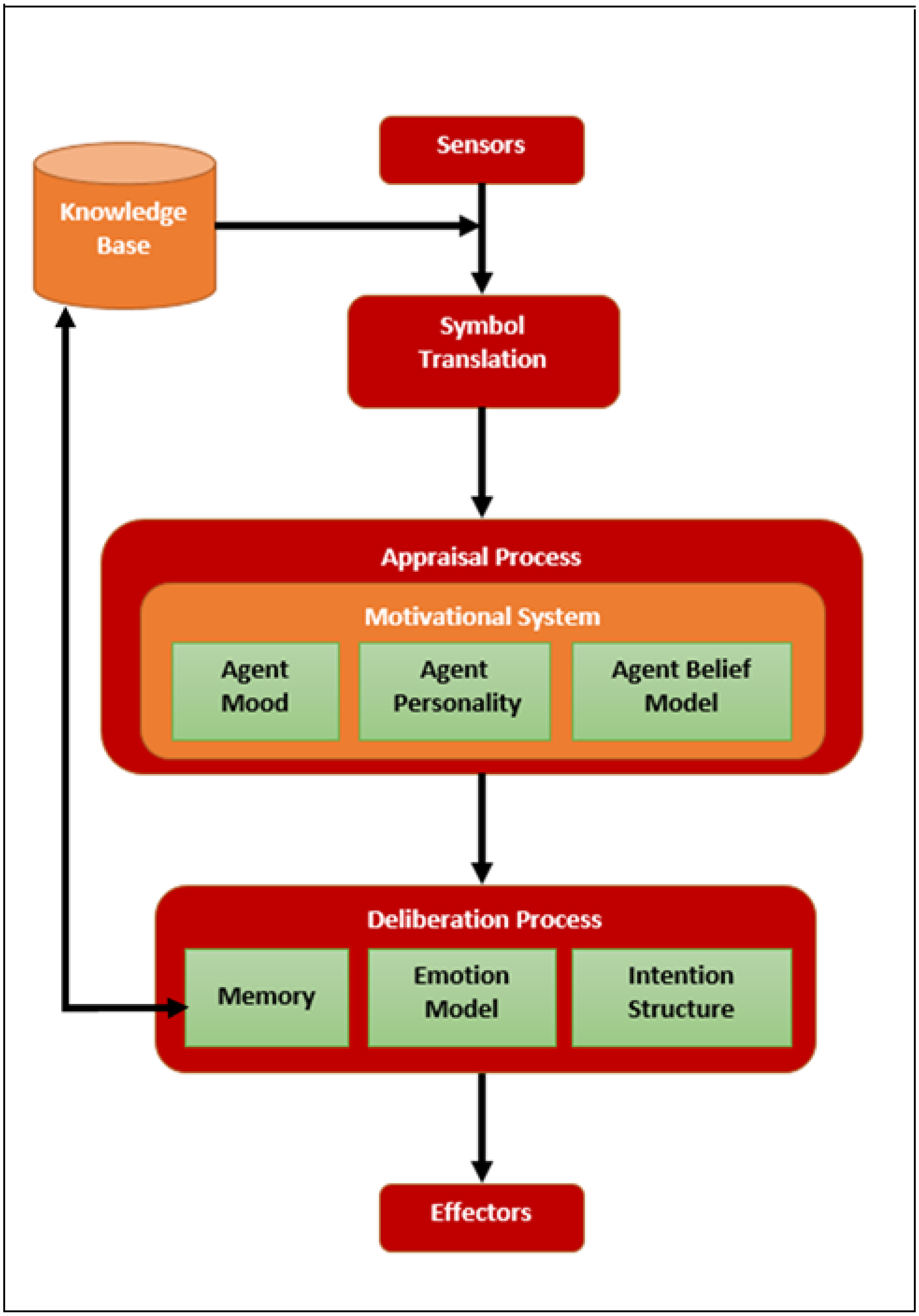


Figure 15: Culturally-Modified Agent Architecture (CMAA)

5.2.1 Knowledge Base & Sensors

- The Knowledge Base is responsible for storing semantic knowledge, such as properties about the world and relations, and the autobiographic memory which contains information concerning past events in the agent's personal experience. We refer to this semantic knowledge as B(Belief) - D(Desire) - (I)Intention framework. Where Belief refers to agent knowledge about the task environment, Desire refers to the goals that the agent wants to achieve within the task environment and the Intention refers to the set of actions that the agent can perform within the task environment (physical simulation of agent behavior).

- Agents perceive the outside world based on their Sensors. The Sensor is responsible for capturing the input and the event from the working environment. For example, within a learning tool input refers to a set of questions queried to the agent, and event refers to the task that needs to be performed within the tool.

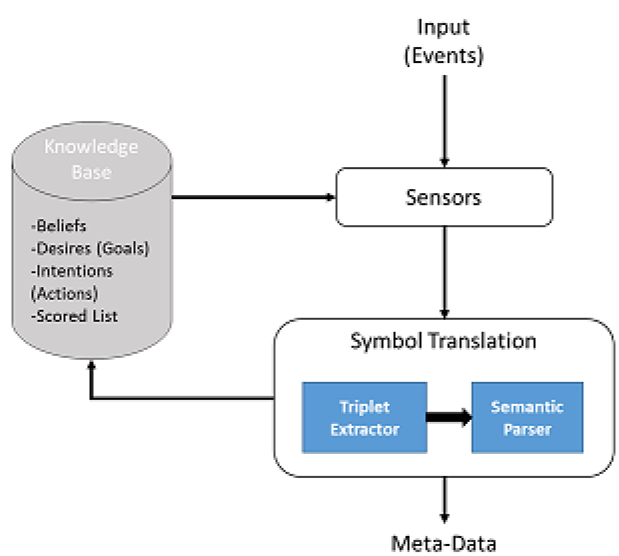


Figure 16: Knowledge base, sensor & symbol translation

5.2.2 Symbol Translation

The Symbol Translation is responsible for encoding the perceived events, and environment knowledge into a set of computationally represented meta-data which acts as the input to the Appraisal Process (see Figure 16).This meta-data is computed using the following NLP tools and techniques:

1. Triplet Extraction: Responsible for extracting subject-action-object triplets from English sentences. A triplet in a sentence is a relation between subject and object, the relation being the action. The aim here is to extract sets of the form subject, action, object out of syntactically parsed sentences using the Stanford Parser [96]. A sentence (S) is represented by the parser as a tree having three children: a noun phrase (NP), a verbal phrase (VP) and the full stop (.). Firstly the subject is found by performing a breadth first search and selecting the first descendant of NP that is a noun or a preposition. Nouns are found in the following subtrees: NN (noun, common, singular), NNP (noun, proper, singular), NNPS (noun, proper, plural) and NNS (noun, common, plural). Secondly, for determining the action (predicate) of the sentence, a search is performed within the VP subtree. The deepest verb descendant of the verb phrase gives the second element of the triplet. Verbs are found in the following subtrees: VB (verb, base), VBD (verb, past tense), VBG (verb, present participle or gerund), VBN (verb, past participle), VBP (verb, present tense), and VBZ (verb, present tense). Thirdly, to find the objects a search is performed in three different subtrees (i.e. all siblings of the VP subtree containing the predicate). The subtrees are: PP (prepositional phrase), NP and ADJP (adjective phrase). Within the NP and PP, the object will be the first noun, while within ADJP it's the first adjective found. Adjectives are found in the following subtrees: JJ (adjective),

JJR (adjective, comparative), JJS (adjective, superlative). Figure 17 shows the working of the Triplet Extraction Algorithm for the following input sentence, ‘I was driving my truck during our regular patrol in Iraq’.

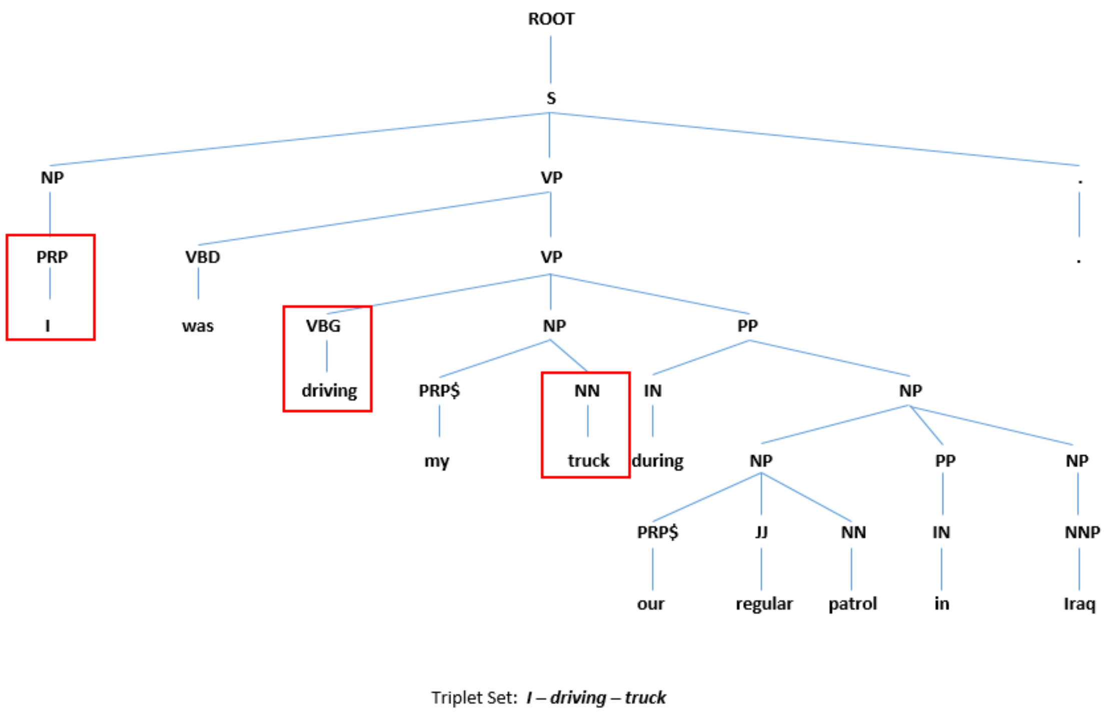


Figure 17: Triplet extraction example

2. Semantic Parser: The triplets extracted from the sentence using the Triplet Extraction Algorithm, act as the input to the semantic parser. The semantic parser was developed to extract attributes (modifiers) and dependencies for each element composing the triplet (subject-action-object). The attributes are responsible for giving more information about the event (e.g. time, place) and the dependency indicates the mutual dependency between the triplets (e.g. and, but, to). The attributes are always found in the modifiers associated with elements, example, the attributes of a noun are mainly adjectives, the attributes of a verb are adverbs. For example, for the input sentence, ‘I was driving my

truck during our regular patrol in Iraq and there was a huge explosion', the output tuples of the semantic parser are shown in Figure 18.

```
Triplet 1: [[subject:'I', subjDep:'None', subjAttr:[], subjPrep:[]],
            [action:'drive', actionDep:'AND', actionAttr:[], actionPrep:[('DURING', 'patrol'), ('IN', 'Iraq')]],
            [object:'truck', objectDep:'None', objectAttr:[], objectPrep:[]],
            [tense:past]]
Triplet 2: [subject:'', subjDep:'None', subjAttr:[], subjPrep:[]],
            [action:'was', actionDep:'None', actionAttr:[('ADV', 'explosion')], actionPrep:[]],
            [object:'truck', objectDep:'None', objectAttr:[], objectPrep:[]],
            [tense:past]]
```

Figure 18: Output from semantic parser

3. Scored List: A scored list of verbs, adjectives and adverbs is maintained in our knowledge database. SentiWordNet 3.0 [97] is used to get the total positive and negative sense count for each verb, adjective and adverb. For example, for the word 'attack', SentiWordNet 3.0 outputs six senses of a verb, and of these senses, one may consider five senses as negative and one as positive. Hence in the knowledge database, the verbs, adjectives and adverbs are stored in the following format: word [Positive Sense Count, Negative Sense Count, Prior Valence]. Equation 2 shows how a prior valence (in range of -5 to +5) is calculated for each selected word. The notion of 'prior valence' also referred as 'semantic orientation' (SO) [98], refers to a real number measure of the positive or negative sentiment expressed by a word or phrase.

$$P(w) = \frac{Tp(w) - Tn(w)}{Ts(w)} * 5.0 \tag{2}$$

Where, P(w) = Prior Valence of word w, whereby -5.0 $\leq$ P(w) $\leq$ +5.0

Tp(w) = Total positive sense count of word w

Tn(w) = Total negative sense count of word w

Ts(w) = Total sense count of word w

A scored list of nouns is also maintained within the knowledge database. To calculate the prior valence of a noun, ConceptNet 5 [99] is used. ConceptNet is a large semantic network of commonsense knowledge which encompasses the spatial, physical, social, temporal, and psychological aspects of everyday life. A value (in range of -5 to +5) is assigned as the valence to an input noun or concept (here we use noun and concept synonymously). To assign valence to a concept, the system collects all concepts which are semantically connected to other concepts and from other concepts to the input concept found in the ConceptNet. The returned entries are separated into two groups depending on their semantic relations. The entries of the first group correspond to relationships like 'IsA', 'DefinedAs', 'MadeOf','PartOf', etc, and the second group entries corresponds to relations like, 'CapableOf', 'UsedFor', 'CapableOfReceivingAction', etc. Of the two groups, the first one basically indicates other associated concepts, and the second one indicates the actions that the input concept can either perform or receive. The first list is searched against the scored list of nouns and the first 5 unique concepts which are found in the target list are taken from the matching list. An average score of those matched 5 concepts is returned as the valence of the non-scored concept.

For example, for the noun 'doctor', the system initially failed to find a prior valence in the existing scored list of nouns. Here, the following two lists are obtained by applying the explained procedures and ConceptNet.

Possible_concept_list = ['person', 'smart person', 'human', 'conscious being', ‘man’, ‘wiley bandicoot’, ‘clever person’, ‘dentist’, ‘pediatrician’, ‘surgeon’, ‘physician’, ‘veterinarian’, ‘messy handwriting’, ‘study medicine’, ‘job’]

Possible_action_list = [‘examine’, ‘help’, ‘look’, ‘examine patient’, ‘help sick person’, ‘wear’, ‘prescribe medicine’, ‘treat’, ‘prescribe’, ‘wear white coat’, ‘look at chart’, ‘save life’, ‘heal person’, ‘take care’] (the list is truncated due to space limitations)

In this case the system first processed the Possible_concept_list, and failed to assign a value. Therefore, the second list, Possible_action_list, is processed and from that list the system returned the value 4.21 by averaging the scores of the verbs, examine (4.50); help (5.00); wear (2.57); prescribe (4.27) and treat (4.69). Hence the value 4.21 is assigned as the prior valence for the concept doctor and stored in the knowledge database for future use.

Just in case, SentiWordNet fails to assign a prior valence to verbs, adjectives and adverbs using SentiWordNet, the similar method of ConceptNet used to calculate its valence value.

The extracted triplets, semantic parser and scored list are sent as input to the Appraisal Process.

5.2.3 Appraisal Process

Once the agent has identified the event, the appraisal process is triggered. In the appraisal process the situations are interpreted, so as to enable a valenced reaction (see Figure 19). The appraisal process consists of the Motivational System:

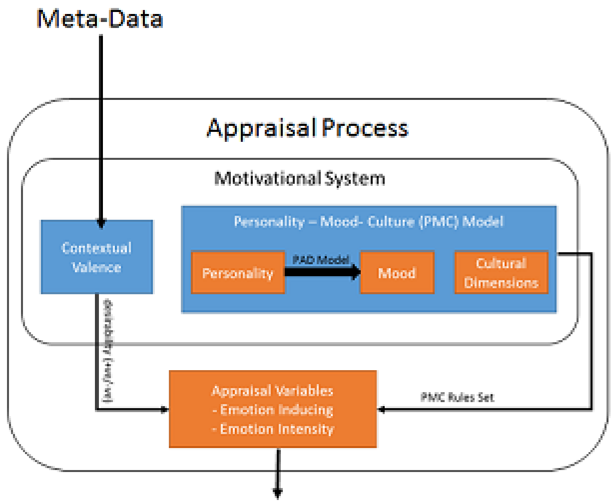


Figure 19: Appraisal process

5.2.3.1 Motivational System

It calculates the desirability of an event depending upon the agent's needs and drives. If an event is perceived as positive, the desirability of the event is set high (i.e. event is perceived as desirable) and if an event is perceived as negative, then the desirability of the event is set to low (i.e. event is perceived as undesirable). The desirability of an event is calculated using the following contextual valence algorithm:

1. Contextual Valence: Before describing the method of Contextual Valence Algorithm, I discuss the concept of Sentiment Sensing or Opinion Mining from text. The words sentiment (e.g., good or bad) and opinion (e.g., positive or negative) are used synonymously with reference to textual data, and the task of sentiment sensing precedes the task of affect or emotion (e.g., happy, sad, anger, hope, etc.) sensing using the emotion model discussed in next step. Various conceptual models, computational methods, techniques, and tools exist for Sentiment Sensing from text.

One such approach is Keyword Spotting and Lexical Affinity that consists of a tagging system that splits the sentence into words and checks through the tagged dictionary to find each word and the corresponding tag category. If a word is not found in the dictionary, the engine will undergo a suffix and prefix analysis. By examining the suffix and prefix of the word, the assumed tag may be derived. The output of the tagging system includes the words and corresponding affect categories. According to a linguistic survey [100], only 4% of the words used in written texts carry affective content. This finding shows that targeting affective lexicons is not sufficient to recognize affective information from texts. Another approach for Sentiment Analysis is using fuzzy logic that assesses an input text by spotting regular verbs and adjectives, without processing their semantic relationships. Here, the verbs and adjectives have preassigned affective categories, centrality, and intensity. As with lexical affinity-based approaches, this method cannot adequately analyze smaller text units such as sentences. Various machine-learning techniques also exist that [101][102] typically rely on affective clues in analyzing a corpus of texts. This approach works well when a large amount of training data of a specific domain of interest (e.g., movie reviews) is available. Although some machine-learning methods identify words and phrases that signal subjectivity, machine learning methods usually assign predictive value to obvious affect keywords. Therefore these are not suitable for sentence-level emotion classification as they fail to incorporate emotion-annotation for other non-affective lexical elements which may have affective connotation. Moreover, machine-learning-based approaches also fail to incorporate rule-driven semantic processing of the words (e.g., contextual valence) used in a sentence.

As my agent data is based on textual case-studies, it was important that the agent architecture breaks down this textual data into sentences and understand the semantic relationship between these sentences. Here we consider each sentence as an independent event that affects the agent's behavior. Hence it was important to implement a Sentiment Sensing technique with the ability to analyze smaller textual units, like sentences which the aforementioned techniques or approaches fail to do. As a result I chose the Contextual Valence Algorithm by M.A.M Shaikh [95], due to its ability of calculating the sentiment of a text by breaking it into smallest possible textual unit and assigning each textual unit its own valence value. I have modified the Algorithm to suit our agent architecture as explained below:

Input: We assume the input is a Paragraph P, containing n sentences, such that $P = \{S_1, S_2,..,S_i,.., S_n\}$ and $1 \leq i \leq n$. As a sentence S_i may have one or more verbs, the semantic parser may output one or more triplet(s) for S_i. We represent S_i as a set of m triplets T, i.e., S_i=$\{T_1, T_2,..,T_j,.., T_m\}$, whereby $1 \leq j \leq m$. A triplet T_j has the following form: *actor, action, object*. The triplet elements actor, action and object have the following form, *name, dependency, prep, attribute*.

Algorithm: The following is a pseudo-code of the Contextual Valence Algorithm underlying our system:

Begin

for each S_i in P do //assume $1 \leq i \leq n$

tripletSet$_i$ = getSemanticParsing (S_i)

//the output of Semantic Parser is a set of Triplets for each sentence.

for each triplet T_j, in tripletSet$_i$ do //we assume $1 \leq j \leq m$, m triplets

actorValence = ContextualValenceAttrib (actorPriorValence, actorAttributes)

actionValence =ContextualValenceAttrib (actionPriorValence, actionAttributes)

objectValence = ContextualValenceAttrib (objectPriorValence, objectAttributes)

actionObjectPairValence=setActionObjectPairVal (actionValence, objectValence)

tripletValence = checkNegation(actorAttributes, actionAttributes, objectAttributes) * setTripletValence (actorValence, actionObjectPairValence)

tripletDependency = if the token "dependency" is found then 'true' else 'false'

tripletDependencyType = if the dependency is "to" then set 'to_dependency' else 'not_to_dependency'

tripletResult$_j$ = tripletValence, tripletDependency, tripletDependencyType

loop until all triplets are processed

contextualValence = processTripletLevelContextualValence (tripletSet$_i$)

sentimentScore = average($\sum_{k=1}^{m}$ abs(contextualValence$_k$))

valenceSign = get ResultantValenceSign(contextualValence)

SentenceValence$_i$ = sentimentScore * valenceSign loop until all sentences are processed

valence = getParagraphValence (SentenceValence)

outputValence = valence $\cup$ SentenceValence

End

The contextual valence of the attributes for actor, action and object are calculated based on a set of rules for Nouns, Verbs, Adverbs, and Adjectives. The algorithm checks for negation and dependency (and, but, to) within a sentence. Once a valence value is attributed to each triplet, the processTripletLevel algorithm is responsible for synthesizing these valence values based on the dependency that exist between the triplets. Finally a sentiment score is assigned to a sentence by averaging the valence values of all triplets and assigning a positive or negative sign. For example, for the input sentence, 'I was driving my truck during our regular patrol in Iraq and there was a huge explosion', the Contextual Valence Algorithm assigns a sentiment score of -4.888 to sentence. Hence the output from the Contextual Valence Algorithm calculates the desirability of the event from the agent's perspective.

5.2.3.2 Personality-Mood-Cultural (P-M-C) Model

Desirability of an event ties into the goals and expectations of the agent. Once the desirability of an event is calculated, the P-M-C (Personality-Mood-Cultural) model is activated to determine the set of values for the appraisal variables necessary for emotion generation. Before explaining the PMC model, lets define the terms that affect the appraisal variables:

1. Mood: Moods reflect medium-term affect, which is generally not related with a concrete event, action or object. Moods are longer lasting stable affective states, which have a great influence on humans cognitive functions [89]. Accordingly the conditions for mood changes can be divided into (a) the onset of a mildly positive or negative event, (b) the offset of an emotion-inducing event, (c) the recollection or imagining of emotional experience, and (d) the inhibition of an emotional response in the presence of an emotion-inducing event. These mood changes are computed within the CMAA across the three dimensions- pleasure (P), arousal (A) and dominance (D) as defined by Mehrabian's PAD Temperament Model [89].

2. Personality: Personality reflects long-term affect. Personality reflects individual differences in mental characteristics. A common personality schema is the Big Five model of personality [8][9]. CMAA computes personality across the five dimensions of the Five Factor Model (FFM): Extraversion (extrovert vs. introvert), Agreeableness (cooperative vs. antagonistic), Conscientiousness (self-discipline), Neuroticism (anger, depression) and Openness (learn, appreciate and experience new things) [8][9].

3. Cultural Belief: Agent's belief (psychological dimensions of an individual) are encapsulated using the Hofstede's Cultural Dimensions [29]. Hofstede's cultural dimensions were chosen because of its ability to analyze and factorize the world and its surrounding events from an agent's perspective across the five dimensions of- Power Distance Index (PDI), Individual- ism/Collectivism (IDV),

Masculinity/Feminity (MAS), Uncertainty Avoidance Index (UAI), Long-term Orientation (LTO):

- Power Distance Index (PDI): Determines the inequality in terms of social status and wealth in society. Power distance is the tendency of an individual to accept that more powerful individuals should have more resources. In CMAA, PDI for an agent was measured in terms of how they react to authority figures, by assigning a negative or positive value (in range -1 to +1). A negative PDI indicates that the agent does not approve of an action associated with an authority figure and a positive PDI indicates that the agent approves of the action associated with the authority figure.
- Individualism/Collectivism (IDV): Determines the ability of an individual to compromise and work in a group environment. In CMAA, IDV for an agent is defined with respect to his in-group or out-group relationships as per the utility of the event. In CMAA, IDV has a negative and positive value associated with it (in range -1 to +1). A negative IDV indicates that the agent believes in working alone (individualistic in nature) and a positive IDV indicates that the agent believes in working for group or believes that decisions should be made with respect to group vs. an individual.
- Masculinity/Feminity (MAS): Refers to the distribution of roles between the genders. In CMAA, MAS for an agent is measured with respect to the agent's interaction with a male vs. a female user. A positive or a negative value (in range -1 to +1) is assigned to MAS based on how an agent view's

the role of male vs. female within a decision making situation or with respect to an authority figure.

- Uncertainty Avoidance Index (UAI): Determines the ability of an individual to bear uncertainty and to make decision without concrete information. In CMAA, UAI is captured by considering agent's reaction to uncertain situations/events that can affect or alter an agent's lifestyle or goals. The UAI has a positive and a negative value associated with it (in range -1 to +1), where a positive value indicates the agent's ability to handle uncertain situations or events, whereas a negative value indicates agents inability to handle uncertain situation or events that can affect its goal.
- Long-term Orientation (LTO): It can be viewed as the tendency to view a set of decisions together as a policy toward an ultimate goal, rather than looking at each decision in its own right. In CMAA, LTO is captured based on how the current interaction/event affects the overall goal of the agent. In CMAA, LTO has a positive and a negative value associated with it (in range -1 to +1). A positive LTO indicates that the current event and situation tend to have an impact on the future goals, whereas a negative LTO value indicates that the goal remains unaffected due to current situation.

A sliding scale was used to assign a value (in the range of -1 to 1) to each of the five cultural dimensions of PDI, IDV, MAS, UAI, and LTO.

The of CMAA can be characterized by specific rules and interplay with several ap-

Table 3: CMAA appraisal variables

Type	Variable Name	Possible Enumerated Values
Event-based	self_presumption (sp)	desirable, undesirable
	other_presumption (op)	desirable, undesirable
	self_reaction (sr)	pleased, displeased
	prospect (pros)	positive, negative
	status (stat)	unconfirmed, confirmed, disconfirmed
Agent-based	action_of_agent (aoa)	approve, disapprove
	direction_of_emotion (de)	self, other
Object-based	object_fondness (of)	like, dislike
Intensity	agent_mood (mood)	exuberant, dependent, relaxed, docile, bored, disdainful, anxious, hostile
	mood_intensity (mval)	numerical calculated value
	effort_of_action (eoa)	high, low

praisal variables based on above defined terms of mood, personality and Cultural Belief. The following subsection defines the set of appraisal variables used within the CMAA architecture and the PMC model rule set to assign values to these variables.

Appraisal Variables: In CMAA two kinds of appraisal variables have been defined, namely emotion-inducing variables (event-based, agent-based and object-based) and emotion intensity variables. These cognitive variables and their values have been defined based on the underlying emotion model we are using, i.e. the OCC emotion model. The structure of OCC emotion model, categories emotions in terms of events, agent and objects [30], hence these appraisal variables were designed. Table 3 lists the appraisal variables used in CMAA along with their possible enumerated values. The event-based variables are calculated with respect to the desirability of the event and cultural beliefs, the agent-based and object-based variables are calculated based on cultural beliefs, whereas the intensity variables are calculated based on the PAD model of personality and mood.

Assigning Values to the Appraisal Variables: This subsection explains the appraisal variables listed in Table 3 and explains the process of assigning enumerated values to those variables using the aforementioned resources.

Table 4: Mood octant's of PAD space

+P+A+D **Exuberant**	-P-A-D **Bored**
+P+A-D **Dependent**	-P-A+D **Disdainful**
+P-A+D **Relaxed**	-P+A-D **Anxious**
+P-A-D **Docile**	-P+A+D **Hostile**

- Agent Mood (mood) and Mood Intensity (mval)

Mehrabian [89] describes mood across three traits of pleasure (P), arousal (A), and dominance (D). The three traits are nearly independent, and form a three dimensional mood space. The implementation of the PAD mood space uses axes ranging from -1.0 to 1.0 for each dimension. Mood is described with the following classification for each of the three mood space axis: +P and -P for pleasant and unpleasant, +A and A for aroused and unaroused, and +D and D for dominant and submissive. With this classification all octant's of the PAD mood space are described by Table 4. If, for example, the mood of a person has the values: 0.25 pleasure, -0.18 arousal, 0.12 dominance, its discrete mood description is slightly relaxed, where a mood represents a point in the PAD space. In CMAA, to calculate the mood of an agent and it's associated intensity the five dimensions of the Five-Factor Model are mapped onto the three dimensional mood space of pleasure, arousal and dominance using the PAD model [89]. Using this mapping (see equations3, 4 and 5), CMAA is able to compute a default mood for the agent:

$$Pleasure = 0.21 * Extraversion + 0.59 * Agreeableness + 0.19 * Neuroticism \quad (3)$$

$$Arousal = 0.15 * Openness + 0.30 * Agreeableness - 0.57 * Neuroticism \quad (4)$$

$$\begin{aligned} Dominance = 0.25 * Openness + 0.17 * Conscientiousness \\ +0.60 * Extraversion - 0.32 * Agreeableness \end{aligned} \tag{5}$$

Using this mapping a person, whose personality is defined with the following big five personality traits: openness=0.4, conscientiousness=0.8, extraversion=0.6, agreeableness=0.3, and neuroticism=0.4 has the default mood slightly relaxed (pleasure=0.38, arousal=-0.08, dominance=0.50). Hence the resultant mood space from the PAD octant will be the enumerated assigned to the agent_mood (mood) variable. Intensity of the mood (mval) is calculated by squaring and summing the values of pleasure, arousal and dominance. The value stored in the mval appraisal variable will help us determine intensity of the emotion generated (explained in next section) which indirectly ties in with the personality and current mood of the agent.

- Self Reaction (sr)

  According to the appraisal structure the values assigned to self_reaction (sr) are 'pleased' and 'displeased' respectively. This variable is assessed with respect to the event taking place. Here the notion of 'self' refers to the agent itself that assess the sentiment of the event as 'positive' or 'negative' from the 'self' perspective. To assign a value to self_reaction (sr), CMAA considers the desirability of the event calculated in the earlier subsection using the Contextual Valence Algorithm. If the desirability of the event (output from Contextual Valence Algorithm) is positive, then we assign the value 'pleased' to (sr), and if the desirability of the event s negative, then (sr) is assigned 'displeased'.

Rule 1:

If sentimentScore ≥ 0:

sr = 'pleased'

If sentimentScore < 0:

sr = 'displeased'

- Self Presumption (sp)

Like self_reaction (sr), this variables is also assessed with respect to the event taking from the agent 'self' perspective. Hence the value assigned to self_presumption (sp) is 'desirable' and 'undesirable'. There are two cases that affect the value of this variable- 1) the desirability of the event and 2) the cultural dimension of Uncertainty Avoidance Index (UAI). To assign the value for appraisal variable (sp), CMAA assumes the following rules: 1) if desirability of an event is positive and UAI is positive the (sp) is set to be 'desirable' for the agent, 2) if desirability of an event is positive and UAI is negative the (sp) is set to be 'desirable' for the agent, 3) if desirability of an event is negative and UAI is positive the (sp) is set to be 'desirable' for the agent, and 4) if desirability of an event is negative and UAI is negative the (sp) is set to be 'undesirable' for the agent.

Rule 2:

If sentimentScore ≥ 0 and UAI ≥ 0:

sp = 'desirable'

If sentimentScore ≥ 0 and UAI < 0:

sp = 'desirable'

If sentimentScore < 0 and UAI ≥ 0:

sp = 'desirable'

If sentimentScore < 0 and UAI < 0:

sp = 'undesirable'

- Other Presumption (op)

  As above, the values for other_presumption (op) are set to 'desirable' or 'undesirable' while assessing the event from the perspective of the agent pertaining to an event being assessed. Here the valenced agent refers to the actorValence calculated using Contextual Valence Algorithm. To assign the values to (op), the following simple rules apply:

  - If a positive valenced event (desirability of event is positive) is associated with a positive valenced agent, (op) is set to 'desirable'. For example, "The teacher was awarded the best-teacher awards". Here the agent is 'teacher' and has a positive valence associated with it (i.e. +4.167) and the event 'award teacher best-teacher awards' has a positive desirability (i.e. +8.741) hence the variable (op) is set to 'desirable' with respect to the agent 'teacher'.

  - If a positive valenced event (desirability of event is positive) is associated with a negative valenced agent, (op) is set to 'undesirable'. For example, "The kidnapper freed the hostage", a negative valenced actor 'kidnapper' (i.e. -4.095) is associated with a positive valenced event (i.e. +5.03) and hence the (op) is set to 'undesirable' for the 'kidnapper'.

- If a negative valenced event (desirability of event is negative) is associated with a positive valenced agent, (op) is set to 'undesirable'. For example, "The teacher punished the student for cheating", a positve valenced actor 'teacher' (i.e. +4.167) is associated with a negative valenced event (i.e. -7.981) and hence the (op) is set to 'undesirable' for the 'teacher'.
- If a negative valenced event (desirability of event is negative) is associated with a negative valenced agent, (op) is set to 'desirable'. For example, "A terrorist escaped from the jail", a negative valenced actor 'terrorist' (i.e. -3.620) is associated with a negative valenced event (i.e. -6.715) and hence the (op) is set to 'desirable' for the 'terrorist'.

Rule 3:

If sentimentScore ≥ 0 and actorValence ≥ 0:

op = 'desirable'

If sentimentScore ≥ 0 and actorValence < 0:

op = 'undesirable'

If sentimentScore < 0 and actorValence ≥ 0:

op = 'undesirable'

If sentimentScore < 0 and actorValence < 0:

op = 'desirable'

- Prospect (pros)

The prospect of an event involves a conscious expectation that it will occur in the future, and the value for the variable prospect (pros) can be either 'positive' or 'negative'. To assign a value to the (pros) variable, CMAA considers the cultural belief of Long Term Orientation (LTO) where (pros) is assigned 'positive' if LTO is positive indicating current event has a positive impact on agent's final goal. Whereas a negative LTO results in (pros) being 'negative', indicating that current situation bears a negative impact on the agent's final goal.

Rule 4:

If LTO ≥ 0:

  pros = 'positive'

If LTO < 0:

  pros = 'negative'

- Status (stat)

  The variable status (stat) has values such as 'unconfirmed', 'confirmed' and 'disconfirmed' associated with it. If the tense of the verb associated with the event is present or future or modal, the value of (stat) is set to 'unconfirmed' for the event. If the verb of the event has positive valence or a confirmation is present in the event (like 'yes') and the tense of the verb is past (like 'I succeeded') then (stat) is set to 'confirmed'. Again, if the verb of the event has negative valence or a negation and the tense of the verb is past with a negation, (stat) is set to 'disconfirmed'.

  Rule 5:

  If eventTense = present, future, modal:

stat = 'unconfirmed'

If verbTense = past and (verbValence ≥ 0 or isAffirmation(event)):

stat = 'confirmed'

If verbTense = past and (verbValence < 0 or isNegation(event)):

stat = 'disconfirmed'

- Direction of Emotion (de)

  Depending on whether the agent that experiences some emotion is reacting to consequences of events for itself or to the consequences for others, the system sets the value of variable direction_of_emotion (de) as either 'self' or 'other'. If a first person pronoun (I, me, my, myself, mine) exists within the event then value of (de) is set as 'self', else the value is set 'other'.

  Rule 6:

  If firstPersonPronounExist(event) = true:

  de = 'self'

  If firstPersonPronounExist(event) = false:

  de = 'other'

- Action of Agent (aoa) and Effort of Action (eoa)

  The value of action_of_agent (aoa) can be either 'approve' or 'disapprove'. The value of (aoa) is based on the cultural dimenion of PDI, where the action of the agent 'self' or 'other' is either praiseworthy or blameworthy based on the event taking place. If (de) is 'self', then the agent is appraising his own actions based on the desirability of the event, i.e. if desirability of the event is negative, the value of (aoa) is set to 'disapprove' and if the desirability of the event is positive, the value of (aoa) is set to 'approve'. If (de) is 'other', then we look at the value assigned to PDI, i.e. if the PDI is negative then the value of (aoa) is set as 'disapprove' and if PDI is positive then the value of (aoa) is set as 'approve'.

Rule 7:

If de = self and sentimentScore ≥ 0:

    aoa = 'approve'

If de = self and sentimentScore < 0:

    aoa = 'disapprove'

If de = other and PDI ≥ 0:

    aoa = 'approve'

If de = other and PDI < 0:

    aoa = 'disapprove'

The variable effort_of_action (eoa) has values 'high' and 'low' assigned to it based on the effort invested, i.e. the greater the effort invested by individual or group the more intense the emotion. Here I compute the value of (eoa) based on the cultral dimension of collectivism (IDV) using the following set of rules: 1) If IDV is negative and (de) is 'self' then (eoa) is set to 'high', 2) If IDV is positive and (de) is 'self' then (eoa) is set to 'low', 3) If IDV is negative, (de) is 'other' and (aoa) is 'approve' then (eoa) is set to 'low', 4) If IDV is negative, (de) is 'other' and (aoa) is 'disapprove' then (eoa) is set to 'high', 5) If IDV is positive, (de) is ‘other’ and (aoa) is ‘approve’ then (eoa) is set to ‘high’ and 6) If IDV is positive, (de) is ‘other’ and (aoa) is ‘disapprove’ then (eoa) is set to ‘low’.

Rule 8:

If de = self and IDV < 0:

eoa = ‘high’

If de = self and IDV ≥ 0:

eoa = ‘low’

If de = other and IDV < 0 and aoa = approve:

eoa = ‘low’

If de = other and IDV < 0 and aoa = disapprove:

eoa = ‘high’

If de = other and IDV ≥ 0 and aoa = approve:

eoa = ‘high’

If de = other and IDV ≥ 0 and aoa = disapprove:

eoa = ‘low’

- Object Appealing (oa)

  The value of object_appealing (oa) indicates whether an interacting agent (here object) is ‘liked’ or ‘dislike’. In order to assign a value to (oa) for an object, CMAA considers the cultural dimension of Masculinity (MAS) and the gender of the interacting agent. If MAS is negative and gender is female then the value of (oa) is set to ‘dislike’, else the value of (oa) is set to ‘like’.

Rule 9:

If MAS $<$ 0 and gender = female:

oa = ‘dislike’

If MAS $<$ 0 and gender = male:

oa = ‘like’

If MAS $\geq$ 0:

oa = ‘like’

Once the appraisal variables are set, the Deliberation process is activated to determine a set of emotions defined by the OCC model and based on the values assigned to these appraisal variables.

5.2.4 Deliberation Process

The Deliberation Process is responsible for activating the emotion model to generate a set of emotions based on the appraisal variables. This then activates the Intention Structure to determine an appropriate action that needs to be performed by the agent, while making an update on the memory (Knowledge Base) of the agent (see Figure 20).

5.2.4.1 Emotion Model

The OCC model of Emotions [30] is used to determine a set of emotional states across each event. The core motivation for choosing the OCC model is that it defines emotions as a valence reaction to events, agents, and objects, and considers valence reactions as a means to differentiate between emotions and non-emotions [30]. The model categorizes emotions into two categories- positive and negative, based on the

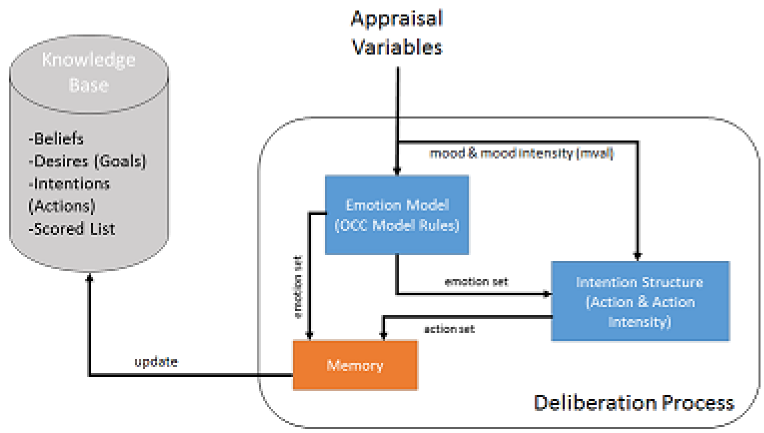


Figure 20: Deliberation process

agents reaction to events, actions, and object within the interacting environment (see Table 5). Emotion recognition consists of inferring a set of emotions by applying a

Table 5: OCC emotion Set

Positive Emotions	**Negative Emotions**
Happy-for	Resentment
Gloating	Pity
Joy	Distress
Pride	Shame
Admiration	Reproach
Love	Hate
Hope	Fear
Satisfaction	Fears-confirmed
Relief	Disappointment
Gratification	Remorse
Gratitude	Anger

set of rules. Depending on whether states expressed by certain appraisal variables (see Table 3) hold or do not hold, multiple emotions can be inferred from a given situation; i.e. the appraisal variables of one rule antecedent can be a proper subset of the antecedent of another rule (see Table 6).

Table 6: Definition of the rules for the OCC emotion types

Emotion	Definition
Joy	Pleased about a desirable event
Distress	Displeased about an undesirable event
Happy-for	Pleased about an event desirable for a liked agent
Pity	Pleased about an event undesirable for a liked agent
Resentment	Displeased about an event desirable for another agent who is not liked
Gloating	Pleased about an event undesirable for another agent who is not liked
Hope	Pleased about positive prospect of a desirable unconfirmed event
Fear	Displeased about negative prospect of a undesirable unconfirmed event
Satisfaction	Pleased about confirmation of positive prospect of a desirable event
Fears-confirmed	Displeased about confirmation of negative prospect of an undesirable event
Relief	Pleased about disconfirmation of negative prospect of an undesirable event
Disappointment	Displeased about disconfirmation of positive prospect of a desirable event
Distress	Displeased about unexpected undesirable event
Joy	Pleased about unexpected desirable event
Pride	Pleased for praiseworthy action/event of self
Shame	Displeased for blameworthy action/event of self
Admiration	Pleased for praiseworthy action/event of other
Reproach	Displeased for blameworthy action/event of other
Gratification	Joy and pride
Remorse	Distress and shame
Gratitude	Joy and admiration
Anger	Distress and reproach
Love	Liking an attractive entity (object or agent)
Hate	Disliking an unattractive entity (object or agent)

Based on the description of the OCC emotions, I have defined the following set of rules assuming an event taking place within a textual data:

Rule 1: Events - Consequences for others

if SR == ‘pleased’ and OP == ‘desirable’:

‘happy-for’ is true

if SR == ‘pleased’ and OP == ‘undesirable’ and AoA == ‘disapprove’ and DE == ‘other’:

‘gloating’ is true

if SR == ‘displeased’ and OP == ‘desirable’ and AoA == ‘disapprove’ and DE == ‘other’:

‘resentment’ is true

if SR == ‘displeased’ and OP == ‘undesirable’ and AoA == ‘approve’ and DE == ‘other’:

‘pity’ is true

Rule 2: Events - Prospect Based

if pros == ‘negative’ and stat == ‘unconfirmed’ and DE == ‘self’ and SR == ‘displeased’ and SP == ‘undesirable’:

‘fear’ is true

if pros == ‘positive’ and stat == ‘unconfirmed’ and SR == ‘pleased’ and SP == ‘desirable’:

‘hope’ is true

if stat == ‘confirmed’ and (‘hope’ is true or pros == ‘positive’):

‘satisfaction’ is true

if stat == ‘confirmed’ and (‘fear’ is true or pros == ‘negative’):

‘fears-confirmed’ is true

if stat == ‘disconfirmed’ and (‘fear’ is true or pros == ‘negative’):

‘relief’ is true

if stat == ‘disconfirmed’ and (‘hope’ is true or pros == ‘positive’):

‘disappointed’ is true

Rule 3: Events: Well-Being

if SP == ‘desirable’ and SR == ‘pleased’:

‘joy’ is true

if SP == ‘undesirable’ and SR == ‘displeased’:

‘distress’ is true

Rule 4: Action of the Agents

if AoA == ‘approve’ and DE == ‘self’ and SR == ‘pleased’ and SP == ‘desirable’:

‘pride’ is true

if AoA == ‘disapprove’ and DE == ‘self’ and SR == ‘displeased’ and SP == ‘undesirable’:

‘shame’ is true

if AoA == ‘approve’ and DE == ‘other’ and SR == ‘pleased’ and OP == ‘desirable’:

‘admiration’ is true

if AoA == ‘disapprove’ and DE == ‘other’ and SR == ‘displeased’ and OP == ‘undesirable’:

‘reproach’ is true

Rule 5: Aspects of Objects

if OF == ‘like’:

‘love’ is true

if OF == ‘dislike’:

‘hate’ is true

Rule 6: Complex Emotions

if ‘joy’ is true and ‘pride’ is true:

‘gratification’ is true if ‘distress’ is true and ‘shame’ is true:

‘remorse’ is true

if ‘joy’ is true and ‘admiration’ is true:

‘gratitude’ is true

if ‘distress’ is true and ‘reproach’ is true:

‘anger’ is true

To show a working example, for the following event:‘I was driving my truck during our regular patrol in Iraq and there was a huge explosion. We were about 10 miles from our base. All I remember is being thrown away on the windsheild and when I woke up I was in the hospital’. the deliberation process generates following set of emotions: [resentment, fear, fears-confirmed, remorse].

Values of appraisal variables mood, mood_intensity (mval), along with the emotion set are then passed on to the Intention Structure to determine the intensity of the emotion being portrayed (low, mid, high) and the appropriate set of actions that need to be performed. The resulting emotion set is also updated to the memory of the agent.

5.2.4.2 Intention Structure

The appraisal variables of mood, mood_intensity (mval) along with the emotion set computed by the OCC emotion model above, act as input to determine the appropriate action (emotion - facial expressions and gesture) and the intensity of the emotion (LOW, MID, HIGH) to be portrayed by the agent based on the following rules:

Rule 1: MID = +ve mood & -ve emotion

Rule 2: HIGH = +ve mood & +ve emotion

Rule 3: MID = -ve mood & +ve emotion

Rule 4: LOW = -ve mood & -ve emotion

These emotion intensity are used to vary the intensity of the facial expressions to be portrayed by the agent, i.e they affect the switch intensity of mouth, brows, eyes and energy of an emotion. We have 22 set of emotions defined for the agent architecture and a typical emotion function would look like this:

```
function sad_emotion() {
    SendText("SetSwitchIntensity [switch= expMouthHappy f0= 0.000000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= expMouthSad f0= 0.690000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= expBrowsSad f0= 0.690000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= expMouthMad f0= 0.000000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= expBrowsMad f0= 0.000000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= expEyesTrust f0= 1.000000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= antiTrust f0= 0.900000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= expEyesDistrust f0= 0.713000 t= 0.4]");
```

```
    SendText("SetSwitchIntensity [switch= antiDistrust f0= 0.358000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= blinks f0= 0.751500 t= 0.4]");
    SendText("SetSwitchIntensity [switch= expBrowsCurious f0= 0.300000 t= 0.4]");
    SendText("SetSwitch [switch= ego state= lessLo]");
    SendText("SetSwitch [switch= agressMaster state= confLess]");
    SendText("SetSwitchIntensity [switch= energyHigh f0= 0.250000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= energyLow f0= 0.000000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= talkBob f0= 0.475000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= headEvadeHighE f0= 0.250000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= headEvadeLowE f0= 0.750000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= HighEnergyNoise f0= 0.250000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= LowEnergyNoise f0= 0.000000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= browTalk f0= 0.250000 t= 0.4]");
    SendText("SetSwitchIntensity [switch= visemes f0= 1.100000 t= 0.4]");
}
```

Hence for the event that is taking place within the agent environment, the resulting emotion set and emotion intensities are updated to the agent memory in the Knowledge Base.

5.2.5 Effectors

Once an emotion set and emotion intensity is generated, an appropriate action (facial expressions, and gestures) is chosen from the agent Knowledge Base. These effectors are responsible for physical simulation of the agent, i.e. generating a verbal response and performing the non-verbal behavior.

5.3 Next Steps

To test the feasibility and effectiveness of our agent architecture CMAA, a VH named Justin was created. Justin was developed as a Virtual Standardized Patient (VSP) portraying an OEF (Operation Enduring Freedom)/OIF (Operation Iraqi Freedom)/OND (Operation New Dawn) Veteran, who has screened positive for the initial screening of mild TBI (Traumatic Brain Injury). In the following chapter I describe the current installation of Justin, our VSP and its interactive system, as a diagnostic tool for clinical evaluation of mild TBI.

CHAPTER 6: VIRTUAL STANDARDIZED PATIENT(VSP) WITH MILD TBI (TRAUMATIC BRAIN INJURY)

This chapter describes the implementation of a Virtual Standardized Patient (VSP) Justin, portraying symptoms of mild TBI (Traumatic Brain Injury), as a diagnostic training tool to: 1) To test the validity and believability of the VH prototype as veteran portraying symptoms of mild TBI and 2) test the effectiveness (conversational effects) of using the VH as a training tool to practice diagnostic evaluation and improve communication between a patient and a provider in a clinical setting. This chapter describes the implications of mild TBI (Traumatic Brain Injury) and how it is measured and evaluated. This chapter also describes the need for a diagnostic tool to practice effective TBI screening communication and standardize the screening process across all veteran facilities.

6.1 Preliminary Work

6.1.1 What is Traumatic Brain Injury (TBI)

Traumatic Brain Injury commonly referred to as TBI is becoming the common cause of death and disability amongst war veterans [22]. Over the last several years, traumatic brain injury (TBI) has been thrust into the forefront of the medical community of the Veterans Office across the country. In the United States each year, an estimated 1.7 million people sustain TBI (Traumatic Brain Injury): 1.365 million are treated and released from an emergency department, 275,000 are hospitalized and

52,000 die as a result of their injuries [22].

The Department of Defense (DoD) and the Department of the Veterans Affairs (VA), have defined TBI as any traumatically induced structural injury and/or physiological disruption of brain function as a result of an external force [22]. TBI injury is graded on three levels mild, moderate or severe based on the level of consciousness or Glasgow coma scale (GCS) score. Mild TBI (GCS 13-15) is in most cases a concussion and there can be a full neurological recovery, although most of these patients suffer from short-term memory and concentration difficulties [21]. In moderate traumatic brain injury (GCS 9-13) the patient is lethargic or stuporous, and in severe injury (GCS 3-8) the patient is comatose, unable to open his or her eyes or follow commands.

6.1.2 Evaluation of TBI Screening Techniques

In a typical TBI screening process, personnel injured during deployment have to submit a self-report (22 item checklist) providing subjective ratings across three domains: concentration, memory and thinking/organization. If screened positive on the self-report, an in-depth evaluation is performed which consists of an interview conducted between the patient and the clinician. Standard measures of TBI screening evaluation (e.g. The Glasgow Coma Scale, American Academy of Neurology Diagnostic Criteria) provide severity ratings, rather than definitive determination of the presence or absence of TBI. Furthermore, the post concussive symptoms are not specific to TBI; many are common in the general population or may be primarily due to other disorders, such as depression, PTSD, or other affective disturbances (see Figure 21). Several studies have looked at the sensitivity and specificity of Veterans Affairs (VA)

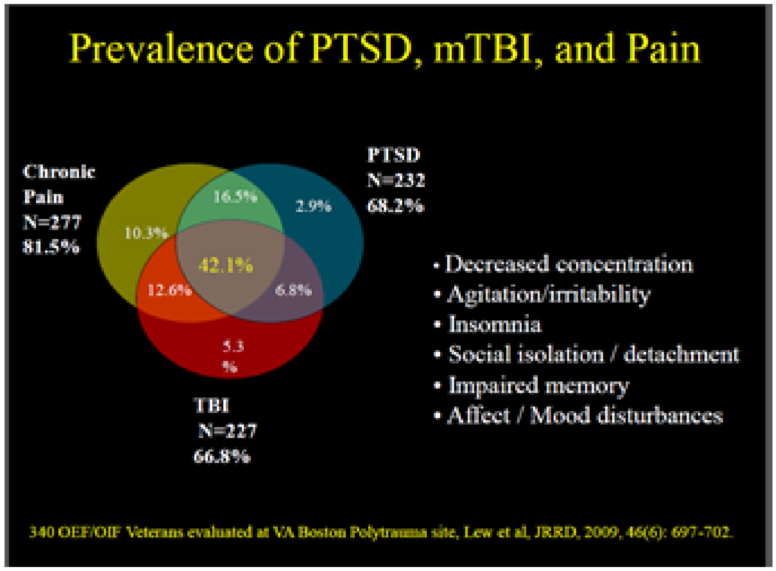


Figure 21: Prevalence of brain injury at VA Polytrauma site [21]

post deployment TBI screening measures. Results of these reports indicate good sensitivity (0.85) when the TBI screening and evaluation were administered on the same day of injury but low sensitivity (0.48) when the screening was administered after lapse of considerable time as part of VA clinical care [19] [20]. Discrepancies were also seen in the screening test conducted by researchers vs. VA clinicians. One study found moderate to high retest reliability over a two-week period when the screening was administered by the researchers [19]. Studies also indicate low-test retest reliability between VA clinician administered TBI screenings vs. research administered TBI screening [20].

6.2 Problem Definition & Research Objective

Due to these differences in the mild TBI screening techniques, the VA's office recognized a need for- 1) provider training to ensure that providers have the required expertise, and 2) a standardized mild TBI screening process across all VA facilities.

Thus a Virtual Standardized Patient (VSP) portraying symptoms of mild TBI, as a diagnostic training tool was identified to practice effective communication of mild TBI screening and standardize the evaluation process across all VA facilities.

6.3 Virtual Standardized Patient (VSP) Portraying Mild TBI

In the first installation, a Virtual Standardized Patient (VSP) portraying an OEF (Operation Enduring Freedom)/ OIF (Operation Iraqi Freedom)/ OND (Operation New Dawn) veteran with symptoms of mild TBI (Traumatic Brain Injury) was developed [23]. The implementation of the VSP included the following objectives: 1) To develop a full description (history, present situation, symptom presentation) of a combat veteran with mild TBI. 2) Design the physical appearance of the VSP and animate it with set of symptoms based on the background description. 3) Implement an interaction between the VSP and the clinician based on the provided sketch. 4) Implement an interaction between the VSP and the clinician based on the CMAA agent architecture to make it more adaptable.

6.3.1 Clinical Background of VSP

The Clinical Team consisting of Ahmed, Bomberger, Cifu, Hurley, Scholten, Taber (based on observation of the comprehensive TBI evaluation (CTBIE) conducted at the Salisbury VAMC site by the Polytrauma Care Team (PCT)) was tasked with creating a case description and history (combat experiences, important events while deployed, all important aspects of present and recent past including major and minor problems and symptoms, family, work, education, etc.) of a recently returned combat Veteran with a history of mild TBI [23]. A full description of a recently returned combat Veteran with mild TBI that includes a detailed history with specific focus on loss/alteration in consciousness was provided by the Clinical team. Additional emphasis was given on outlining the traumatic experiences and events, as well as delineation of current and recent issues relating to family, work, education, and symptom presentation (verbal and nonverbal) in multiple domains. A descriptive report on TBI was provided (167 pages), from which information on behavioral and cognitive characteristics and symptoms manifestations of a typical mild TBI patient was extracted (see Table 7).

Table 7: TBI behavioral charecteristics

Domain	Symptom	Visual Correlate
Physical	Headache	Squinting, rubbing head
	Dizziness	Blinking, tilting head
	Fatigue	Dark circles, red eyes,red rims around eyes, rubbing eyes, nodding off
	Sensitivity to light/noise	Squinting, wears sunglassesindoors
	Blurred vision	Blinking, Rubbing eyes
Cognitive	Attention Difficulties/ Difficulty concentrating	Breaks in eye contact,increased movement, increased self-touching (itching, rubbing head,picking at nails)
	Memory deficits	Furrowing brow, appearing to have forgotten, looking down,touching forehead
	Speed of processing(slowed)	See Behavioral Correlate
	Impaired Judgment	See Behavioral Correlate
Behavioral Correlate	Depression	Decreased hygiene, lack of expression, downward gaze,sighing, tearfulness
	Anxiety	Restless, darting eyes,more rapid speech, sweating, fidgets
	Impulsivity	Blurting things out, touching things on desk
	Irritability	Appears annoyed

From the clinical case study provided, information about patient's family background, medical history, deployment history, and injury symptoms was extracted and populated within our patient database. A set of questions and answers observed during a typical CTBIE, and informative commentary was also provided by the Clinical team, which was populated (around 10-20 pages of textual data) within our Knowledge Base.

6.3.2 VSP Development

A Caucasian male patient (age 30) named Justin was designed and built, based on the physical description provided to us by the VA office (see Figure 22). Justin was animated (see Figure 24) with set of animations (like headache, anger, sadness, and irritation) based on the behavior characteristics provided to us in the case description (see Table 7). Haptek software was used to develop the physical appearance of the VSP and the animation library for the facial expressions and gestures associated with the symptoms and emotions portrayed by the VSP.

6.3.3 VSP Interaction Interface

A web-based training tool was developed consisting of two parts: 1) a life-sized version of VSP projected on one dedicated screen (see Figure 22), and 2) an interaction interface (see Figure 23) for the users to interact with the VSP on an one-on-one basis. A typical interaction between the user and the VSP consists of the user asking the VSP questions (text or speech) based on the interview and evaluation guidelines outline by the VA system for CTBIE. The VSP perceives these questions and generates an appropriate verbal feedback, and an emotional state (portrayed using

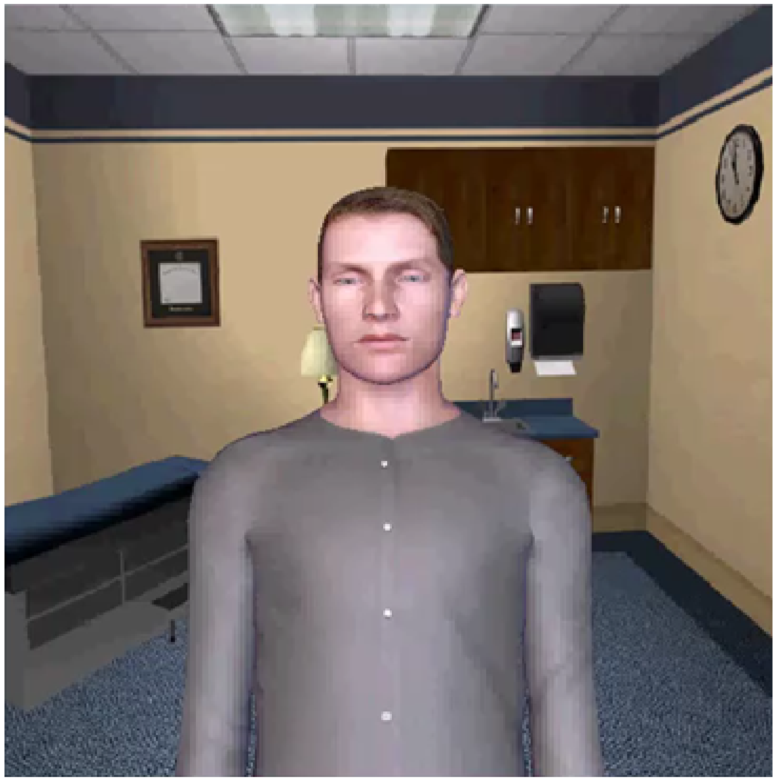


Figure 22: Virtual patient with mild TBI

symptoms, facial expressions and gestures) based on the information stored within the VSP's knowledge base. In both the versions of the Interface, the verbal response was generated using the underlying information retrieval algorithm.

6.3.4 Version I of VSP

In the first version of the interaction, the VSP was modeled to portray symptoms and set of basic emotions like- anger, happiness, sadness, fear, shyness, pondering, and skeptic. These emotions are emoted by changing VSPs facial expressions (eyes squinting, direction of eyebrows, tilt of mouth, etc.) and posture (hunched back in sadness, chest thrust forward in anger, crouching in fear, folded arms or hand on chin to ponder, etc.), (see Figure 24). These set of emotions and symptoms portrayed by the VSP were scripted by our experts based on their observations of a typical

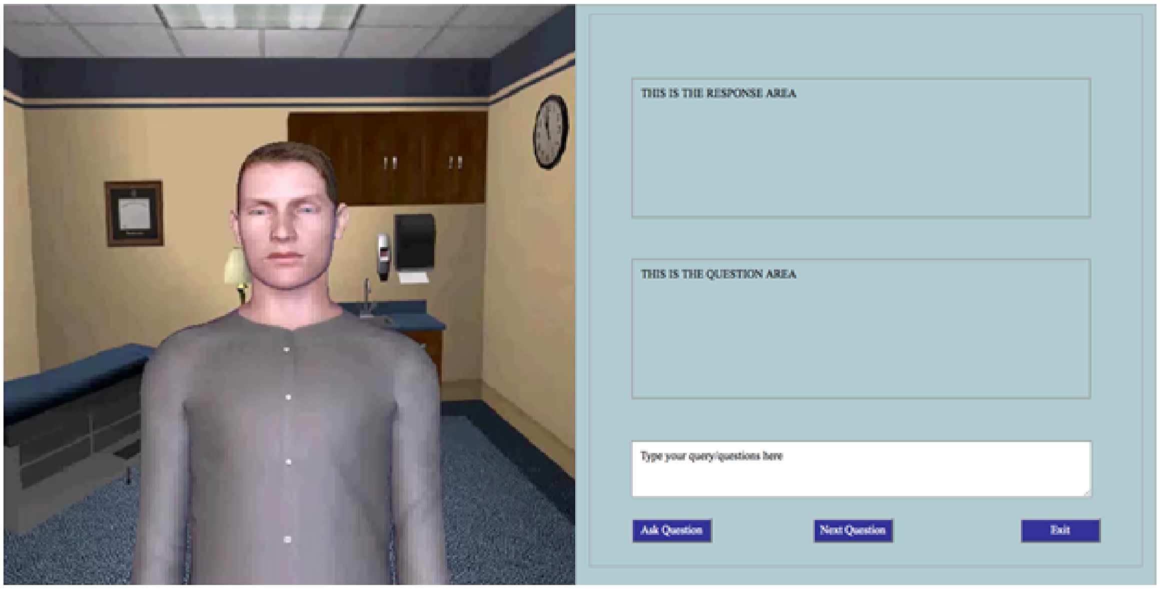


Figure 23: Web-based interaction interface with VSP

TBI screening between a veteran and a clinician at the polytrauma clinic. A sample interaction (see Table 8) between the user and VP is shown populated with emotions and non-verbal behavior:

Table 8: Sample interaction with emotions and gestures

Person	Verbal Output	Emotion	Non-Verbal Output
User	Good Morning, Mr. Smith. How are you doing today?		listening gesture (head nod, folded arms, eye contact)
VP	Good morning Doctor. I am not feeling so great today.Can you tell me more about todays session?	sad	Random sad expression (hunched back, averted gaze, sad nod, arms folded in sadness, mouth tilted downwards, drooping lower eyelids)
User	In this session, I will be asking you more detailed questions so that we can better understand your symptoms and be able to come up with specific diagnosis. Do you have any questions at this time?		listening gesture (head nod, folded arms, eye contact)
VP	So the initial screening didnt give any indication about any form of trauma to my brain. And todays second screening will help you determine if I am suffering from any Traumatic Brain Injury?	fear	Random fear expression (mouth wide open, raised eyebrows, neck thrust back)
User	Yes that is correct. We will be performing a more detailed evaluation today.		listening gesture (head nod, folded arms, eye contact)
VP	Okay	calm	Random neutral expressions and talking gesture (hand movements)
User	Can you describe the events surrounding your accident in Iraq or Afghanistan?		listening gesture (head nod, folded arms, eye contact)
VP	I was driving my truck during our regular patrol in Iraq and then there was a huge explosion. We were about 10 miles from our base. All I remember is being thrown away on the windshield. It was bad, very bad doctor.	fear	Random fear expression (mouth wide open, raised eyebrows, neck thrust back)
User	Did you lose consciousness?		listening gesture (head nod, folded arms, eye contact) Symptoms: Blinking, tilting head
VP	Yes, I think I lost consciousness for few minutes. But I am not too sure about it.	Skeptic, pondering	Hand on chin, head bowed in thought, one eye and eyebrow drooped Symptoms: Blinking, Rubbing eyes
User	After the event, how was your ability to function? Was it affected at all?		listening gesture (head nod, folded arms, eye contact)
VP	It was very bad doctor. I could not continue to perform my duty as before. It was very difficult. I could not focus and was very slow in doing day-to-day things. I knew that something was wrong with me.	Sad, Agitated	Random sad expression (hunched back, averted gaze, sad nod, arms folded in sadness,mouth tilted downwards, drooping lower eyelids) Symptom: Rubbing head, rubbing eyes, eyes darting

Figure 24: Emotions and expressions of Virtual Standardized Patient

6.3.5 Version II of VSP

In the second version of the VSP interface, the cognitive behavior of the VSP was driven by the underlying CMAA agent architecture, explained in the earlier chapter. The semantic knowledge about our VSP is determined from the Case Study provided to us by the Clinical experts, given short sample:

"Mr. Vet is a 30 year old Caucasian male, Operation Iraqi Freedom (OIF)/Operation Enduring Freedom (OEF) Veteran who presented to the OIF/OEF intake clinic at his local VAMC. He screened positive for TBI and was referred for a comprehensive TBI evaluation (CTBIE) with the Polytrauma Care Team (PCT).

Previous records are suggestive of a possible TBI from exposure to a close proximity IED explosion during his deployment two years ago. There was no loss of conscious-

ness (LOC) but he did report feeling confused for several minutes with "time moving in slow motio". He sustained superficial lacerations on his face and upper chest. He was taken to the base clinic where physical exam revealed that he was alert, oriented to name, place and date, but somewhat slow in responses. Physical and neurological exam was unremarkable. His superficial wounds were cleaned and required only conservative management. He was cleared to resume duty the following day.

Mr. Vet reported that he started to experience some new symptoms 2-3 days following the event including headaches (HA) and difficulty in focusing and paying attention. He described 1-2 episodes of HA every week lasting for about an hour. Pain is 6/10 in intensity and located in the right temple without any radiation or aura. He reported sensitivity to light and noise and a short nap appeared to be helpful. The HA responded well to standard dose ibuprofen. The symptoms progressed with time. His attention span decreased and he noted difficulty with prolonged reading. He compensated with constant reminders and shorter tasks during his deployment. Mr. Vet was honorably discharged 6 months after his return to the United States.

In the PCT clinic, Mr. Vet complains of physical and cognitive symptoms; HA, generalized pain, sleep disturbances, short attention span, inability to focus, memory issues, and unable to plan and carry out simple tasks. Behaviorally, he reported significant mood-swings, verbal aggression which was directed towards his wife and kids, and depression. He denied any SI/HI (suicidal/homicidal ideation). He is isolating himself from friends and has withdrawn from activities that he previously used to enjoy. Functionally, he has difficulty with completion of complex tasks such as managing his finances and participating in college classes. Overall, his physical, cognitive, and *behavioral issues have significantly impacted his day to day functioning and adversely affected his marriage and family relationships. His physical exam was normal and his neurological exam did not show any focal deficits. He recently completed brain MRI which was normal. His labs were in normal range.*

Based on the symptoms at the time of injury, he was diagnosed with a history of mild TBI, with associated depression, headaches, insomnia and cognitive deficits. He was started on Amitriptyline for headache and sleep and set to follow-up routinely at the PCT clinic. He has also engaged with the Mental Health team which has initiated Zoloft and individual counseling for treatment of PTSD."

This case study is a textual document, hence information extraction tecniques were used to filter out important inofrmation pertaining to deplyment history of the veteran, like the number of years of active duty,where the veteran was deployed over the years, injury reports of all injuries sustained during deplyoment, symptoms and behavior report of the patient, and so on. This sematic knowledge is then stored within the Knowledge Base of the VSP. The Knowlede Base also incorporated the the genral behavioral charecteristics associated with patients suffering from mild TBI, as shown in Table 7. The semantic knowledge represents the (B)elief of the VSP, the behavioral characteristics of mild TBI patients represnts the (I)ntention of the VSP, and the goal of recieving optimum care and determining the presence of mild TBI represents the (D)esire of the VSP.

The event percieved by the Sensor of the CMAA agent architecture is represented by the set of questions being asked to the VSP within our interaction interface. As mentioned earlier, these questions are based on the interview and evaluation guidelines outline by the VA system for CTBIE. These questions are then passed onto the Appraisal Process where the desirability(+ve/-ve) of the event is calculated based on the semantic knowledge of the VSP. The appriasal variables are then computed by mapping the desirabilty of the event along with P-M-C (Personlaity-Mood-Culture) set values. These appriables variables form intput to the motion model that generates a range of emotions (there are 22 different emotions based on the OCC model) for that one particulare event of the VSP. These emotions alson with the intensity variables of the aprriasal process are then passed to the Intention Structure to choose an appropriate action of the VSP. The Intention Structure is also responsible to determine the set of symptoms portrayed by the VSP. Here NLP techniques are used to extract lexicons (from the verbal feedback) associated with the symptoms described within the TBI behavior table. These actions (emotions) and symptoms are then updated into VSP memory (Knoweldge Base) for the particular event taking place. Finally the emotions and the symptoms are then performed by the effectors in terms of non-verbal physical behavior, along with the corresponding verbal feedback for the event.

A sample interaction between the user and VP is shown (see Table 9) based on the emotion generation determined by the CMAA agent architecture:

Table 9: VSP interaction & emotion generation using CMAA agent architecture

Screening Question	Verbal Feedback	Emotion Set
I am going to ask you to describe for me in detail events you experienced in Iraq or Afghanistan where you may have been physically injured. Please focus on the event that was most powerful or injurious and emotionally disturbing ?	I was driving my truck during our regular patrol in Iraq and there was a huge explosion. We were about 10 miles from our base. All I remember is being thrown away on the windsheild and when I woke up I was in the hospital.	resentment fear remorse
Do you recall if it was an IED, RPG, Motor Grenade or a Bomb ?	I did not know at that time. But later I found out that it was an IED explosion that knocked our truck.	resentment fear anger
Did you remember losing consciousness after the incident ?	Yes I think I lost consciousness for a few minutes. But I am not too sure about it. Yes I was. All I remember when I gained conciousness was being carried to our base hospital.	resentment fear anger
Can you recall what really happened with you ? Can you recall the approximate distance from the blast site ? Can you talk about any of your physical injuries ?	I dont clearly remember the incident in details. But like I said before I was driving the patrol car in Iraq and suddenly there was this explosion. I think it was about 10-15 feet from our truck. I cant be very sure as the explosion was very sudden and unexpected. I remember hitting onto the windshield and losing consciousness. When I gained consciousness I was being carried to the base hospital. I was confused for quiet some time and unaware of my surroundings. But I remember I was bleeding and had bruises all over my face and arms and no mobility in my legs.	resentment relief remorse
Did any of the objects penetrate your skull during the explosion ?	The doctors at the base hospital had me x-rayed and they didnt find any fragments or objects in my skull.	resentment fear anger

6.4 Next Steps

To test the effectiveness of using a VSP as a diagnostic tool for clinical evaluation of mild TBI, a usability study was conducted. The following chapter describes our hypothesis, the experimental setup and the results of the usability study.

CHAPTER 7: EVALUATING THE EFFECTIVENESS OF VIRTUAL STANDARDIZED PATIENT (VSP)

7.1 Introduction

A usability study was conducted between the two versions of the Virtual Standardized Patient (VSP) to determine the effectiveness of developing intelligent virtual agents as a task-based learning tool within a clinical setting. In the following sections I will be describing the various studies that were conducted and present the results from the same.

7.2 Evaluation Study

To test the effectiveness of the CMAA agent architecture and the usability of the VSP Interface as a diagnostic tool, the following evaluation studies were conducted:

7.2.1 System Evaluation by Experts

Clinical experts from the polytrauma clinic at NC Salisbury, responsible for sketching and designing the patient data and the case-study based on their observations of a typical TBI screening were recruited to evaluate the physical appearance of the VSP and the interaction interface.

Experimental Setup: Experts were asked to interact (one-on-one basis) with the web-based VSP tool based on the set of questions outlined in a typical mild TBI screening. A typical interaction with the VSP consists of asking a set of questions (in a serial order) defined within the interface using a set of navigation buttons, and observing the response and the behavior of the VSP to each of the question asked. After the interaction the experts were asked to fill out a quantitative and qualitative survey to evaluate the VSP on the following criteria:

- Evaluate the VSP portraying symptoms of mild TBI:
  1. Physical appearance of the VSP
  2. Verbal feedback of the VSP
  3. Non-verbal behavior of the VSP (facial expressions/gestures)
  4. Symptoms portrayed by the VSPt synonymous with symptoms portrayed by a typical TBI patient.
- Evaluate the web-based Interaction tool:
  1. Usability of the web-based interface
  2. Questions & Answers posted and displayed within the Interface
     - Wording and language of the questions and answers being posted
     - Order in which the questions are being posted
  3. Interaction time
     - Response time between question posted and feedback received

Results: Clinical experts (N = 4) experienced in TBI screening from the VA polytrauma clinic at Salisbury, were recruited as the experts to evaluate the VSP and the web-based interface. A typical interaction consisted of the expert using the web-based interface to interact (ask a set of questions) with the VSP. The experts were asked to evaluate the VSP's behavior (i.e. appearance, gestures, emotions and symptoms) and the usability and learnability associated with the web-based tool. The experts were also asked to check the syntax and language of the questions being asked and the answers being generated. After the interaction the experts were provided with an online survey consisting of quantitative and qualitative questions. In the quantitative part, the experts were asked to rate the various aspects of VSP and the interface by providing a score on a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). In the qualitative part of the survey, the experts were asked to provide feedback on the various aspects of the system. In the following subsection I describe the quantitative and qualitative results from the expert evaluation.

- Web-based Interface Evaluation: Overall the quantitative score for the web-based interface tool were quite high ($M = 4.425$, $SD = 0.67$).Experts indicated that the system was easy to use and found it to be well integrated ($M = 4.375$, $SD = 0.75$). Experts found the interaction time between question asked and the response generated to be very fast ($M = 4.5$, $SD = 0.5$). Experts found the order and syntax of the questions to be in a proper format ($M = 4.25$, $SD = 0.83$).

- VSP Evaluation: Overall quantitative scores indicate that the VSP's appearance and behavior matched the description that was provided to us by the experts ($M = 3.79$, $SD = 0.86$). Experts indicated that the appearance of the VSP was realistic and matched that of the sketch provided of a war veteran ($M =$

Table 10: System evaluation by experts (N=4)

Evaluation Criteria	Sub-Topic	N	Mean	Std. Dev	Std. Error
Web-based Interface	Overall	4	4.425	0.67	0.105
	Usability	4	4.375	0.75	0.24
	Integration of Q&A	4	4.25	0.83	0.41
	Interaction Time	4	4.5	0.5	0.25
Virtual Standardized Patient (VSP)	Overall	4	3.79	0.86	0.13
	Physical Appearance	4	4.125	0.86	0.21
	Verbal Feedback	4	4.00	1.00	0.35
	Non-Verbal Behaviour	4	3.87	0.85	0.21
	Emotion Generation	4	4.5	0.5	0.25
	Symptom Generation	4	3.5	0.5	0.14

4.125, SD = 0.59). Scores also indicate that experts found the VSP's response (verbal feedback) to be fast, clear and easy to understand (*M = 4.00, SD = 1.00*). Scores also indicate that experts found the VSP's non-verbal behavior (gestures, facial expressions) to be realistic (*M = 3.87, SD = 0.85*) and the VSP's emotions identifiable (*M= 4.5, SD = 0.5*). Finally, experts indicated that the symptoms were identifiable (*M = 3.5, SD = 0.5*), though we were provided feedback for making the symptoms more effective.

- Qualitative Feedback: Experts comments, feedback and suggestions are summarized as follows: Overall feedback suggested that the experts found the tool easy to learn and use. Experts also found the VSP's appearance to be realistic and emotions easily identifiable and relateable. Though experts had suggestions to improve the set of questions by breaking bigger questions into set of follow-up questions. Experts also provided suggestions to improve the non-verbal behavior of the VSP by generating a set of different actions (gestures) associated with the symptoms. Currently we have fixed set of action associated with each symptom. Experts also indicated that providing a brief description of the VSP (details about the deployment history, period of injury, and daily life after injury) would be helpful during clinical evaluation. These suggestions and changes will be inculcated within the next iteration under the guidance of our clinical team.

7.2.2 Pilot Study

A pilot study was conducted to evaluate the Experimental Setup and Design of the Usability Study. Students from the introductory course of Computer Science were recruited as participants for the pilot study.

Participants: 15 students from the Computer Science Department at the University of North Carolina, at Charlotte participated in the pilot study on a voluntary basis.

Experimental Task: The participants were asked to perform one interaction of the interface that included the following task: pre-experiment questionnaire, testing session, post-condition questionnaire, testing session, post-condition questionnaire, post-experiment questionnaire. Same testing conditions and setup were used to conduct the pilot study as defined in the experimental procedure of the next subsection.

Observations: The pilot study was conducted to evaluate the following factors:

- Total time required for entire usability study, inclusive of the pre and post experimental setup.
- Time required for each testing session.
- Errors observed during interaction with the VSP.
- Errors observed during navigation of the interface.
- Language and Grammatical Syntax of the Questions and Answers of the VSP.

- Language and Grammatical Syntax of the Questionnaires.

Based on the participant feedback and my observations corrections were made to the interface.

7.2.3 Usability Study of the VSP Interface

7.2.3.1 Experimental Setup

- Subjects: 30 Students from the nursing department were recruited for the study as participants. Participants were required to be at least 18 years of age and proficient in English.

- Variables:

  Independent Variables (IV): Version 1 (scripted emotions & behavior generation) vs. Version 2 (automated emotions & behavior generation using CMAA architecture) of the VSP interface.

  Dependent Variables (DV): Change in cognitive behavior of the VSP between the Independent variables- Version 1 (scripted) and Version 2 (non-scripted CMAA).

- Experimental Design: A within group study was conducted with students participating in each of the following condition:

  Condition Scripted: Students get to interact with scripted version (Version 1) of the VSP interface (scripted behavior of the VSP).

  Condition Automated: Students get to interact with the automated version (Version 2) of the VSP interface (behavior generated using the CMAA agent architecture).

- Counterbalancing of the Conditions: During the testing session, the conditions were presented in the same order to all the participants. The interaction across both the condition in terms of symptom generation, order of the questions and the verbal feedback was same. The only difference being in the behvior generation of the VSP. The participants were not told about which interface they are interacting with during the testing session. The only instruction given to them was to observe the non-verbal behavior of the VSP and the verbal feedback and indictae the emotions and the symptom manifestation in a short survey. Hence I didn‘t swap the testing conditons, as I believed that no affect would be observed.

- Hypothesis:

  $\underline{H_0}$: No significant difference observed with the VSP interaction for *Condition Scripted* vs. *Condition Automated*.

  $\underline{H_1}$: No significant difference observed within the behavior portrayed by the VSP for *Condition Scripted* vs. *Condition Automated*.

7.2.3.2 Experimental Procedure

An interaction between the participant and the VSP interface consist of the following procedure (see Appendix B): pre-experiment session, interaction with *Condition Scripted*, post-session questionnaire 1, interaction with *Condition Automated*, post-session questionnaire 2, and the post-experiment questionnaire, amounting to approximately one hour of completion time.

- Pre-Experiment Session: The participants are given a participants information sheet to read and were asked if they had any questions about the study session. Once done, the participants sign a Consent Form (see Appendix B). A pre-experiment questionnaire is then provided to the participants to collect the following information:

  1. Demographic Information of the student
  2. Experience using an online interactive tool for screening.
  3. Knowledge about Traumatic Brain Injury (TBI) or other Mental Health Disease.
  4. Experience interacting with a TBI patient or other Mental Health Disease.

- Testing Session: The participants then get to interact with the VSP interface in each condition (*Scripted* and *Automated*). For each condition participants get to interact on an one-on-one basis with the VSP using a web-based interface tool. For each session participants get to pose questions to the VSP using navigation buttons present within the interface. The questions are text-based and presented in a sequential order. The participants are then asked to observe the verbal and non-verbal response of the VSP. For each question asked, the participants are provided with a short survey questionnaire.

  <u>Survey Questionnaire</u>: For each Q&A posted during the interaction participants were asked to indicate the following information:

1. Symptoms portrayed by the VSP.
2. Emotions generated by the VSP.

- Post-Session Questionnaire: A post-session questionnaire is then provided to the participants at the end of each testing session (see Appendix B), i.e one after *Condition Scripted* and one after *Condition Automated*. The participants are asked to rate the interaction based on the following criteria:

  1. Quantitative evaluation of VSP based on behavior and appearance.
  2. Quantitative evaluation of the VSP interaction.
  3. Qualitative feedback, comments and critiques.

- Post-Experiment Questionnaire: A post-experiment questionnaire is provided to participants at the end of the testing session (see Appendix B). Within the post-experiment questionnaire participants are asked to rate the various aspects of the VSP interface:

  1. System Usability Scale (SUS) to evaluate the usability and learnability of the Interface.
  2. Comparative evaluation of the VSP in terms of differences observed in *Condition Scripted* vs. *Condition Automated* of the interaction.
  3. Quantitative scale to indicate the effectiveness of using a VSP interface as a diagnostic tool for mild TBI screening.
  4. Qualitative feedback, comments and critiques.

7.2.3.3 Experimental Results

30 participants from the School of Nursing at University of North Carolina, at Charlotte participated in the study. Participants got to interact in both *Scripted* and *Automated* condition of the testing session. Participants were recruited by classroom announcement and recruitment letter. The average age of the participants was in the age group 25-34. Participants were required to be at least 18 years of age and able to communicate comfortably in English. Out of a total of 30 participants, 24 participants were female and 6 participants were male. All the participants were seniors in the Family Nurse Practitioner (FNP) program. The role of a FNP is that of a primary care provider, which includes the necessary background education to diagnose and treat TBI and other illness and injury conditions.To provide hands-on experience with clinical evaluations and diagnostic practices, the School of Nursing is equipped with the Simulation Lab, that allow Family Nurse Practitioners to interact with rela-life actors portraying symptom manifestation to practice communication and screening techniques.

Hence results from the study can impact the usefulness (i.e. accessibility and learnability) of a TBI diagnostic tool for training of the FNP's, as comared to real-life actors in the simulated lab. The results from the evaluation study are summarized as follows:

- How accurately were the participants able to guess the emotions and symptoms portrayed by the VSP in *Scripted* and *Automated* condition ?

  During the testing session, for each condition the participants were asked to fill out a survey questionnaire for every Q & A interaction with the VSP. Participants were asked to indicate the emotions and symptoms being portrayed by the VSP during the interaction. For *Scripted* condition, 88.54% of the emotions were guessed accurately and 92.5% of the symptoms were accurately guessed by the participants. For *Automated* condition, 91.25% of the emotions were guessed correctly and 89.12% of symptoms were guessed correctly by the participants. Results indicated that the for emotion like 'Fear' participants tend to mix it up with 'Sadness' a lot.

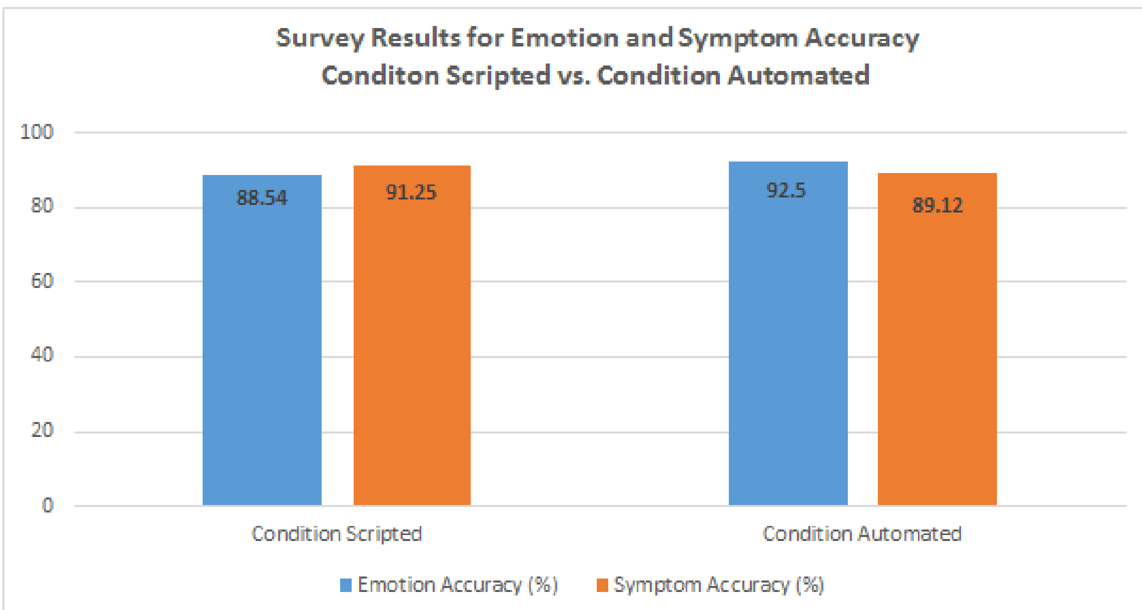


Figure 25: Percent of emotions and symptoms accuracy

- Were differences observed with the VSP interaction and behavior for Condition *Scripted* vs. Condition *Automated* ?

  After each testing session of *Scripted* and *Automated* condition , the participants were provided with a post-session questionnaire to rate the VSP interaction on a Likert scale (1= strongly disagree, 5 = strongly agree). Overall, the VSP in

Table 11: VSP interface evaluation (N=30)

Evaluation Criteria	**Sub-Topic**	**N**	**Mean**	**Std. Dev**	**Std. Error**
Overall VSP Evaluation	Condition Scripted	30	4.16	0.68	0.034
	Condition Automated	30	4.49	0.63	0.030
VSP Interaction Evaluation	Condition Scripted	30	4.25	0.68	0.055
	Condition Automated	30	4.55	0.60	0.049
VSP Behavior Evaluation	Condition Scripted	30	4.12	0.49	0.052
	Condition Automated	30	4.44	0.66	0.049

Automated condition (*M = 4.49, SD = 0.63*) was rated higher than interaction in *Scripted* conditon (*M = 4.16, SD = 0.68*). A t-test was used to compare the differences in VSP interaction and behavior in *Scripted* and *Automated* condition , as rated by the participant. There was a significant difference observed in the VSP interaction between *Scripted* condition (*M = 4.25, SD = 0.68*) and *Automated* condition (*M = 4.55, SD = 0.60*); t(58)=2.00, p = 0.007; thus rejecting our first hypothesis H_0. There was a significant difference observed in the VSP behavior between *Scripted* condition (*M = 4.12, SD = 0.69*) and *Automated* condition (*M = 4.44, SD = 0.66*); t(58)=2.00, p = 0.018; thus rejecting our second hypothesis H_1. Thus results show that the interaction with VSP in *Automated* condition and the behavior portrayed by the VSP in *Automated* condition was significantly better, thus validating that CMAA can create believable intelligent VHS and provide a better interaction environment in a social setting.

- System Usability Scale (SUS):

  The System Usability Scale (SUS) questionnaire, developed by Digital Equipment Corporation, consisting of 10 questions on a Likert scale (1= strongly disagree, 5 = strongly agree), was administered to the participants after the

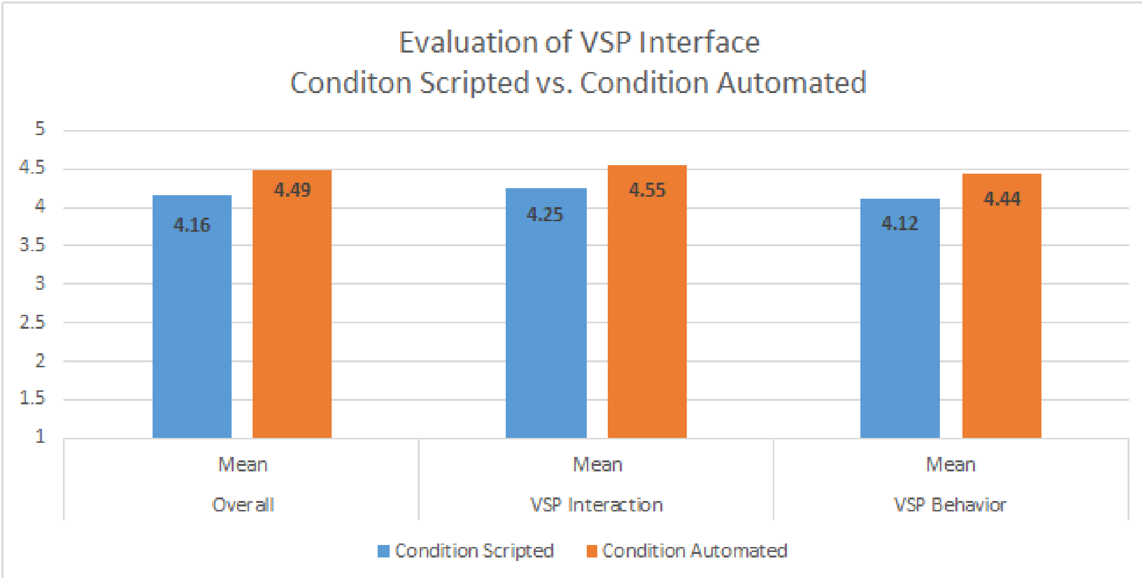


Figure 26: VSP interface evaluation

testing session as part of the post-experiment questionnaire. Overall, the system exhibited a high degree of reported usability (M = *80*, SD = *9.37*) with a highest usability score of 97.5 and lowest score of 62.5.

- User Interface Evaluation Questionnaire:

  A User Interface Evaluation questionnaire, consisting of 8 questions was administered to the participants as part of the post-experiment questionnaire. The participants were asked to evaluate the User Interface based on the Organization and Language of the Content (1 = Confusing, 5 = Very Clear), Positions of the messages and navigational buttons (1 = Inconsistent, 5 = Consistent), and the Learning associated with interface (1 = Difficult, 5 = Easy). Overall, the results for the quantitative evaluation scores were quite high (M = *4.36*, SD = *0.61*) indicating that, the participants found the system to be well integrated and easy to learn.

- Virtual Standardized Patient Evaluation Questionnaire:

  Overall, the results for the quantitative evaluation were quite high (*M = 4.22, SD = 0.69*). Participants indicated that the overall interaction (in terms of non-verbal behavior) with VSP in *Automated* condition was better than that in *Scripted* condition (*M = 4.10, SD = 0.70*). Participants also indicated having VSP portray varying personality and emotions would be more helpful (*M = 4.6, SD = 0.48*) within a social training environment.

  As mentioned earlier our participnats are seniors in the Family Nurse Practioner program. As part of their curriculum, seniors are provided with a simulation lab where they train with real-life actors potraying sympotms of various health issues as a mean of practicing diagnostic evaluation, symptom identification, and practicing screening communication. Hence we wanted to see how the participants would rate the VSP as compared to working with actors in terms of behavior and symptom manifestation portrayed by the VSP.Results from the questionnaire indicated that participnats found interacting with the VSPin both the *Scripted* and *Automated* condition more helpful than using an actor (*M = 4.46, SD = 0.66*).

- Diagnostic Tool Evaluation Questionnaire:

  Overall, the results from the Diagnostic Tool Evaluation questionnaire were quite high (*M = 4.20, SD = 0.65*), indicating that the participants thought the VSP Interface would make a good diagnostic tool within a clinical setting for mild TBI screening. Participants also indicated that having a VSP with varying

Table 12: Post-experiment evaluation (N=30)

Evaluation Criteria	**Sub-Topic**	**N**	**Mean**	**Std. Dev**	**Std. Error**
System Usability Scale (SUS)	Overall	30	80	9.37	1.71
User Interface Evaluation	Overall	30	4.36	0.61	0.039
VSP Evaluation	Overall	30	4.22	0.69	0.048
	VSP Cond *Automated* vs. Cond *Scripted*	30	4.1	0.70	0.057
	VSP Personality & Emotions	30	4.6	0.48	0.089
	VSP better than actor	30	4.46	0.66	0.122
Diagnostic Tool Evaluation	Overall	30	4.20	0.65	0.045
	Cond *Automated* vs. Cond *Scripted*	30	3.83	0.77	0.14

personality and emotions, i.e. Condition 2 would make a better diagnostic tool than a scripted VSP in Condition 1 ($M = 3.83$, $SD = 0.77$). Thus indicating that VSPs portraying symptoms and behavior of mild TBI would make a good diagnostic tool for evaluation of TBI.

- Post-Experiment Qualitative Questionnaire

  Participants were asked to provide qualitative feedback on the following aspects of the VSP interface:

  1. Overall Appearance of the VSP.
  2. Overall Behavior of the VSP.
  3. Interface as a Diagnostic Tool.
  4. Learning associated in terms of TBI.
  5. Overall feedback, suggestions and critique.

  We recieved a total of 450 reponses (each participant was asked a total of 15 qualitative questions). Out the 450 responses, around 80.66

  The most common feedback observed across both the conditions was the VSP's ability to portray minor expressions and show realistic appearance:

“Interesting to see the VSP’s ability to emote using facial expressions.”

“Realistic than expected.”

“Ability to interact and not be still.”

Participants also noticed the subtle variations in the behavior of the VSP in terms of intensity of the emotions potrayed:

“Ability to potray various intensity of facial expressions and gestures was helpful.”

“Was helpful to see emotions and symptoms related to TBI.”

Participants are also noted that the VSP interface would make a better training tool than using actors:

“It is a helpful tool for communication and training practice.”

Participants also noted that they did gain some knowledge regarding TBI, but a little background information before the session would have been helpful:

“I enjoyed this interface and though it has a good potential for nursing students.”

“I did understand the cognitive behavior associated with a TBI patient, but more knowledge about the TBI from military perspective will be helpful.”

Participants indicated that they would like to see the VSP move and also suggested providing deployment history of the patient:

“Make it more movable and give brief background about the patient deployment history.”

Almost all participants indicated that they would use the VSP interface as training tool if it were made available to them:

“I think it could be very helpful as a training tool as it is easily accessible than an actor and helpul to study the different behavior of a patient with mild TBI.”

Though some of the negative comments provided by the participants were in terms of the mobility of the VSP:

“Making the patient walk around the doctor’s office and show some agitated behavior would have been more helpful.”

“I prefer using an actor as opposed to this tool, as it provides a better fidelity in terms of understanding body language of the patient.”

Based on the qualitative feedback, I recognize a need to provide some knid of background about the TBI patients in terms of their case history before the start of the interface. I also plan to animate the VSP with mobility and more symptom manifestation by migrating the VSP to the DoD recognized Virtual Human Toolkit.

7.2.4 Observations of the Evaluation Studies

These are some of the observations of the evaluation study conducted:

1. Counterbalancing of the conditions: In the current user study conducted, the participnats first interacted with *Scripted* condition of the VSP and then the *Automatated* condition. I believe, that swapping the testing conditons (i.e have participnats interact with the *Automatated* condition first then the *Scripted*)

won't have significant difference on the results, as the participnats were not informed about the premises of the interacting interface. Acorss both the conditon of the VSP interface, the set of question, answers and animation asociated with the behavior and symptom generation was kept the same. But I plan to take this into consideration for the our next set of user study.

2. The current evaluation study was conducted by recruiting seniors of Family Nurse Practioner program at the School of Nursing. This tool is being developed as a diagnostic tool to practice TBI screening at the VA facility for clinicians practicing at the VA facility with knowledge about TBI. Hence, the next set of user studies will be conducted at the VA facility, to evaluate the learning outcomes associated with using VSP as a diagnostic training tool.

CHAPTER 8: CONCLUSION & FUTURE WORK

Human to virtual human interaction may be regarded as the next frontier in interface design, particularly for tasks that are social or collaborative in nature. Virtual Human (VH) interfaces while challenging to develop and evaluate, have the potential to revolutionize the accessibility, usability, and applicability of computers in everyday life. Research evidence suggests that people can accomplish tasks more effectively when the behavior and attitude of an VH is similar to a real human. Since virtual humans are modeled after humans, these interface agents can use several modalities for communicating information, such as gesture, facial expressions, which are transparent or obvious to the user. Research in the area of Virtual Human Interfaces is scattered among a variety of fields, including agent systems, animated characters, user emotions, graphics and animation, user and cognitive modeling, conversational interface agents, animated pedagogical agents, virtual reality and human factors.

My primary area of research lies in developing intelligent agent architectures with the ability to plan, make decisions, choose appropriate actions, update memory, display cognitive behavior and generate emotions. Specifically, the goal of my work is to generate automated intelligent agents based on agent's past, education, personal life, personal experiences, cultural beliefs and personality to investigate how these capabilities affect social factors such as engagement, satisfaction, acceptance of the interface agents role, and the success of task performance in an interactive social or public setting. To this end, I have developed an agent architecture called the Culturally Modified Agent Architecture (CMAA) with the ability to portray intelligent behavior based on personality, mood and the cultural belief of the agent. To test the feasibility of CMAA we developed a Virtual Standardized Patient (VSP) portraying symptoms of mild TBI (Traumatic Brain Injury) was developed and implemented it within a clinical setting as a diagnostic tool for training clinicians in TBI screening. The following subsection presents conclusion on each aspect of my research project.

8.1 Dr. Chestr Show

My initial research objective was to develop VHs with the with the ability to interact naturally and spontaneously using speech, personality and emotions. As a result Dr. Chestr a virtual game show host, infused with a unique personality to promote user engagement and enjoyment, was developed. Dr. Chestr was designed to test users with questions about the C++ programming language and allows user to communicate using the most natural form of interaction, speech. To drive the behavior of Dr. Chestr (verbal output and non-verbal output) an agent architecture was developed (see Figure 13) based on natural language understanding, personality and action generation.

A usability study was conducted, to test the effectiveness of Dr. Chestr as a study tool and test if his personality was identifiable. Results from the evaluation study indicated that the users were able to identify with Dr. Chestr's personality. Dr. Chestr was shown to enhance user experience and made the interaction more enjoyable. Though not statistically significant, the increase in users PANAS scores after interacting with Dr. Chestr can have significant implications in the recruitment and retention of computer science students. The successful development of Dr. Chestr with an identifiable personality (extroverted, personable, intelligent, slightly conceited and sarcastic virtual human) provided the foundation for the development of VHs that can portray personality and emotions to enhance user experience and learning outcomes.

8.2 Culturally Modified Agent Architecture (CMAA)

Our preliminary work showed that it is possible to create autonomous Virtual Humans using an intelligent agent architecture with the ability to portray personality, mood and emotions that are adaptable across different domain or application areas. Human emotions are directly influenced by one's culture, attitude, ethics, personality and values. Thus I developed the Culturally Modified Agent Architecture (CMAA) based on 1) agent's belief (past history and personal experiences), 2) agent's personality and 3) mood of the agent. The two main factors of CMAA, which govern an agent's behavior, are: the appraisal process (responsible for appraising events as desirable or undesirable based on mood, personality and agents belief) and the deliberation process (responsible for generating an emotional state and actions). This mapping from the appraisal process to the deliberation process was done using a set of newly defined rules, termed as P-M-C (personality, mood and cultural belief's) rule set. The output from the P-M-C rule model is a set of 22 different emotions and 8 mood factors, which are responsible for varying the intensity of the emotions.

8.3 Virtual Standardized Patient (VSP)

To test the feasibility and effectiveness of our agent architecture CMAA, a VH named Justin was created. Justin acts as a Virtual Standardized Patient (VSP) portraying an OEF (Operation Enduring Freedom)/OIF (Operation Iraqi Freedom)/OND (Operation New Dawn) Veteran, who has screened positive for the initial screening of mild TBI (Traumatic Brain Injury). A case-study of a mild TBI patient was provided by the clinical experts at the VA polytrauma clinic, Salisbury, NC. Based on this case-study Justin, the VSP was developed and a web-based interaction tool was created for training purposes. Two different versions of the VSP interface were developed. In the first version, the cognitive behavior (emotions and symptoms) of the VSP was scripted by the experts. In the 2'nd version, the cognitive behavior of the VSP was automated and driven by the CMAA agent architecture. A typical interaction between the user and the VSP consist of the user asking the VSP questions based on the interview and evaluation guidelines outline by the VA system for screening patients with mild TBI.

8.4 Evaluation of the VSP

A usability study was conducted to evaluate the interaction and behavior of the VSP, and test it's effectiveness as a diagnostic tool; by recruiting students from the School of Nursing at UNC, Charlotte. Participants were asked to interact with both the versions of the VSP interface and provide rating across various aspects like - interaction with the VSP, emotions portrayed by the VSP, symptoms generated by the VSP, usability and learnability of the user interface and effectiveness of VSP as a diagnostic tool. Results showed that a high percentage of emotions and symptoms portrayed by the VSP were accurately guessed by the participants. Results also showed that significant difference was observed by the participants in terms of the interaction and behavior of the VSP in version 1 vs. version 2 of the interface, thus rejecting our null hypothesis. These results indicate that due to its ability to generate varying cognitive behavior the 2'nd version of the VSP interface would make a better diagnostic tool for training and interaction. Hence we can conclude that - 1) It is possible to create autonomous intelligent VHs whose behavior is affected by one's personal history (deployment history of veteran), and personal experiences (behavior of the veteran) and 2) these intelligent VSP's can be used as a training tool as part of simimulation lab, to provide a resource for training nurses or clinicians. Future evaluation studies will be coducted across all VA facilites to validate the use of automated, adaptive, intelligent VSP's as a learning and diagnostic tool to improve interaction and communication within a social task-based training environment.

8.5 Future Work

Cultural differences are one of the main contributors of disparities in the health care industry with respect to the quality of services provided. Smedley et al., found that cultural differences including language and geography resulted in disparate health care received by minorities insured at the same level as the white majority. Resolutely, there is a significant need to include culture and related topics into curricula, especially nursing, to help reduce and eventually eliminate the racial and ethnic health disparities. Research indicates significant existence of racial and ethnic disparities in access to health care service. The Healthy People 2020 initiative, launched by the U.S Department of Health and Human Services, has emphasized the need to eradicate these disparities and thereby improve the health of all groups. Therefore, it has become necessary to provide culturally competent medical care to improve the quality of the health care industry. To provide culturally relevant care, it is necessary to acknowledge patients cultural practices and take their cultural influences into account. A major challenge in the health care industry is to educate and prepare future nurses with skills in trans-cultural nursing.

Building upon the foundations of my doctoral research and a need to incorporate transcultural training within a diagnostic tool, my future plans include the following:

1. Implement the Virtual Standardized Patient (VSP) as part of the simulation lab at the School of Nursing as a diagnostic tool to train nursing students in effective communication and training techniques necessary for primary health care.

2. Investigate the effect of gender and skin color of the VSP on the communication between clinician and VSP in an inter-cultural setting.

3. Investigate the effect of cross-cultural VSP (incorporate culture of different countries to drive the behavior of the VSP) on the communication between clinician and VSP in an inter-cultural setting.

REFERENCES

[1] Badler, Norman I., Diane M. Chi, and Sonu Chopra. “Virtual human animation based on movement observation and cognitive behavior models.” Computer Animation, 1999. Proceedings. IEEE, 1999.

[2] Brooks, Rodney A. “How to build complete creatures rather than isolated cognitive simulators.” Architectures for intelligence (1991): 225-239.

[3] Champandard, Alex J. AI game development: Synthetic creatures with learning and reactive behaviors. New Riders, 2003.

[4] Mascardi, Viviana, Daniela Demergasso, and Davide Ancona. “Languages for Programming BDI-style Agents: an Overview.” WOA. 2005.

[5] Chamberlain SP, Recognizing and responding to cultural differences in the education of culturally and linguistically diverse learners. Intervention in School and Clinic, 2005, 40(4), 195-211.

[6] T. L. Cross, B. J. Bazron, K. W. Dennis et al., Toward a Culturally Competent System of Care, Vol 1: Washington D.C.: National Institute of Mental Health, Child and Adolescent Service System Program (CASSP) Technical Assistance Center, Georgetown University Child Development Center, 1999.

[7] McColIum, C., et al. “Applying a cognitive architecture to control of virtual non-player characters.” Simulation Conference, 2004. Proceedings of the 2004 Winter. Vol. 1. IEEE, 2004.

[8] Digman, J M, 1990. “Personality Structure: Emergence of the Five-Factor Model, Annual Review of Psychology, Vol. 41: 417 -440, DOI:10.1146/annurev.ps.41.020190.002221.

[9] McCrae, R. R. and John, O. P. (1992), An Introduction to the FiveFactor Model and Its Applications. Journal of Personality, 60: 175215. doi: 10.1111/j.1467-6494.1992.tb00970.x.

[10] Jeff Rickel, Stacy Marcella, Jonathan Gratch, Randall Hill, David Traum, William Swartout, “Toward a New Generation of Virtual Humans for Interactive Experiences, IEEE Intelligent Systems, pp. 3238, July/August, 2002.

[11] Jonathan Gratch, Stacy Marsella, “Evaluating the Modeling and Use of Emotion in Virtual Humans, Autonomous Agents and Multiagent Systems, International Joint Conference on, pp. 320-327, Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1 (AAMAS04), 2004.

[12] Mincer, Jacob. ”Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.” (1974).

[13] Skinner, Burrhus Frederic. Science and human behavior. Simon and Schuster, 1953.

[14] B. Smedley, A. Stith, A. Nelson, eds. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington D.C.:National Academy Press, 2002.

[15] Browne, Colette and Alice Broderick "Asian and Pacific Island Elders: Issues for Social Work Practice and Education. Social Work, Vol 39, No 3 (May 1994): 252-259.

[16] National Center for Health Statistics. 1997. National Health Interview Survey. Centers for Disease Control, U.S. Department of Health and Human Services: Washington, D.C.

[17] Healthy People 2010. 2nd ed. Washington, D.C: Department of Health and Human Services, 2000.

[18] Institute of Medicine. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington: National Academics Press; 2002.

[19] Donnelly, Kerry T., James P. Donnelly, Mina Dunnam, Gary C. Warner, C. J. Kittleson, Janet E. Constance, Charles B. Bradshaw, and Michelle Alt. "Reliability, sensitivity, and specificity of the VA traumatic brain injury screening tool." The Journal of head trauma rehabilitation 26, no. 6 (2011): 439-453.

[20] Fortier, Catherine Brawn, Melissa M. Amick, Alexandra Kenna, William P. Milberg, and Regina E. McGlinchey. "Correspondence of the Boston Assessment of Traumatic Brain Injury-Lifetime (BAT-L) Clinical Interview and the VA TBI Screen." The Journal of head trauma rehabilitation (2013).

[21] Rimel RW, Giordani B, Barth JT, et al. Disability caused by mild head injury. Neurosurgery 1981; 9: 322128.

[22] http://www.polytrauma.va.gov/

[23] Mohammed Ahmed, Raghavi Sakpal, Dale-Marie Wilson, David Cifu, Joel Scholten, Chloe A. Bomberger, Robin A. Hurley, Christine Elnitsky, Katherine H. Taber. (2014). Creating a Virtual Mild TBI Patient: Conceptual Design and Initial Database Development, American Neuropsychiatric Association 2014 (ANPA).

[24] Raghavi Sakpal (2012). Virtual patients to teach cultural competency. In Proceedings of the 14th ACM international conference on Multimodal interaction (ICMI '12). ACM, New York, NY, USA, 349-352. DOI=10.1145/2388676.2388754.

[25] Sakpal, R. and Wilson, D.M. & "Can Culture Translate to the Virtual World?", HCI INTERNATIONAL 2011 - Communications in Computer and Information Science, 2011, Volume 173, Part III, 242-246, DOI: 10.1007/978-3-642-22098-2_49.

[26] Sakpal, R, Wilson, D.M. "Virtual game show host Dr. Chestr", Virtual Reality Conference (VR), 2011 IEEE , vol., no., pp.237-238, 19-23 March 2011 DOI: 10.1109/VR.2011.5759486.

[27] Dale-Marie Wilson, and Raghavi Sakpal, Dr. Chestr: Computerized Host Encouraging Students to Review', IVA 2010.

[28] Shaikh, Mostafa Al Masum, Helmut Prendinger, and Mitsuru Ishizuka. "Sentiment assessment of text by analyzing linguistic features and contextual valence assignment." Applied Artificial Intelligence 22, no. 6 (2008): 558-601.

[29] Hofstede, Geert. "Culture and organizations." International Studies of Management & Organization (1980): 15-41.

[30] Ortony, A., Clore, G., Collins, A.: Cognitive Structure of Emotions. Cambridge University Press, Cambridge (1988).

[31] http://www.haptek.com/

[32] Watson, David, Lee A. Clark, and Auke Tellegen. "Development and validation of brief measures of positive and negative affect: the PANAS scales." Journal of personality and social psychology 54.6 (1988): 1063.

[33] Cassell, Justine; Sullivan, Scott; Prevost (2000), Embodied Conversational Agents, Cambridge, MA: MIT Press.

[34] Rickel, Jeff, Stacy Marsella, Jonathan Gratch, Randall Hill, David Traum, and William Swartout. Toward a new generation of virtual humans for interactive experiences. UNIVERSITY OF SOUTHERN CALIFORNIA MARINA DEL REY INFORMATION SCIENCES INST, 2002.

[35] Swartout, William R., Jonathan Gratch, Randall W. Hill Jr, Eduard Hovy, Stacy Marsella, Jeff Rickel, and David Traum. "Toward virtual humans." AI Magazine27, no. 2 (2006): 96.

[36] Imison M, Hughes C (2008) The virtual patient project: using low fidelity, student generated online case studies in medical education. inHello? Where are you in the landscape of educational technology? Proceedings ascilite Melbourne 2008.

[37] Stevens A, Hernandex J, Johnsen K, et al. The Use of Virtual Patients to Teach Medical Students Communication Skills. The 2005 Association for Surgical Education Annual Meeting; 2005 April 710; New York, NY.

[38] Amy Stevens, Jonathan Hernandez, Kyle Johnsen, Robert Dickerson, Andrew Raij, Cyrus Harrison, Meredith DiPietro, Bryan Allen, Richard Ferdig, Sebastian Foti, Jonathan Jackson, Min Shin, Juan Cendan, Robert Watson, Margaret Duerson, Benjamin Lok, Marc Cohen, Peggy Wagner, D. Scott Lind, The use of virtual patients to teach medical students history taking and communication skills, The American Journal of Surgery, Volume 191, Issue 6, June 2006, Pages 806-811, ISSN 0002-9610, DOI: 10.1016/j.amjsurg.2006.03.002.

[39] Kenny, P., Parsons, T., Gratch J., Leuski, A., Rizzo A.: Virtual Patients for Clinical Therapist Skills Training. 7th International Conference on Intelligent Virtual Agents, pp 197-210, Paris France. (2007).

[40] Donald, Merlin. Origins of the modern mind: Three stages in the evolution of culture and cognition. Harvard University Press, 1991.

[41] Baylor, Amy L. "Permutations of control: Cognitive considerations for agent-based learning environments." Journal of interactive learning research 12, no. 4 (2001): 403-425.

[42] Bickmore, Timothy Wallace. "Relational agents: Effecting change through human-computer relationships." PhD diss., Massachusetts Institute of Technology, 2003.

[43] Johnson, W. Lewis. "Interaction tactics for socially intelligent pedagogical agents." In Proceedings of the 8th international conference on Intelligent user interfaces, pp. 251-253. ACM, 2003.

[44] Ryokai, Kimiko, Cati Vaucelle, and Justine Cassell. "Virtual peers as partners in storytelling and literacy learning." Journal of computer assisted learning 19, no. 2 (2003): 195-208.

[45] Plant, Amy L. Baylor E. Ashby. "Pedagogical agents as social models for engineering: The influence of agent appearance on female choice." Artificial intelligence in education: Supporting learning through intelligent and socially informed technology 125 (2005): 65.

[46] Baylor, Amy L., E. Shen, Daniel Warren, and S. Park. "Supporting learners with math anxiety: The impact of pedagogical agent emotional and motivational support." In Workshop on "Social and Emotional Intelligence in Learning Environments," held at the International Conference on Intelligent Tutoring Systems., Macei, Brazil. 2004.

[47] Sanghoon, Amy L. BAYLOR Daniel WARREN, and PARK E. SHEN Roberto. "The impact of frustration-mitigating messages delivered by an interface agent." Artificial intelligence in education: supporting learning through intelligent and socially informed technology 125 (2005): 73.

[48] Baylor, Amy L., and Yanghee Kim. “Simulating instructional roles through pedagogical agents.” International Journal of Artificial Intelligence in Education 15.2 (2005): 95-115

[49] Ryu, J. E. E. H. E. O. N., and Amy L. Baylor. “The psychometric structure of pedagogical agent persona.” TECHNOLOGY INSTRUCTION COGNITION AND LEARNING 2, no. 4 (2005): 291

[50] Nass, Clifford, and Youngme Moon. “Machines and mindlessness: Social responses to computers.” Journal of social issues 56, no. 1 (2000): 81-103

[51] Ulicny, Branislav, and Daniel Thalmann. Crowd simulation for interactive virtual environments and VR training systems. Springer Vienna, 2001.

[52] Takeuchi, Yugo, and Yasuhiro Katagiri. “Social character design for animated agents.” Robot and Human Interaction, 1999. RO-MAN’99. 8th IEEE International Workshop on. IEEE, 1999.

[53] Rickel, Jeff, and W. Lewis Johnson. “STEVE (video session): a pedagogical agent for virtual reality.” Proceedings of the second international conference on Autonomous agents. ACM, 1998.

[54] Isbister, Katherine, and Patrick Doyle. “Design and evaluation of embodied conversational agents: A proposed taxonomy.” The First International Joint Conference on Autonomous Agents & Multi-Agent Systems. 2002.

[55] Isbister, Katherine, and Patrick Doyle. “Design and evaluation of embodied conversational agents: A proposed taxonomy.” The First International Joint Conference on Autonomous Agents & Multi-Agent Systems. 2002.

[56] Evers, Marc, and Anton Nijholt. “Jacob-An animated instruction agent in virtual reality.” Advances in Multimodal InterfacesICMI 2000. Springer Berlin Heidelberg, 2000. 526-533.

[57] Babu, Sabarish, et al. “Can Immersive Virtual Humans Teach Social Conversational Protocols?.” Virtual Reality Conference, 2007. VR’07. IEEE. IEEE, 2007.

[58] Nouri, Elnaz, and David Traum. “A cross-cultural study of playing simple economic games online with humans and virtual humans.” Human-Computer Interaction. Applications and Services. Springer Berlin Heidelberg, 2013. 266-275.

[59] Lin, Yang, and Andrew S. Rancer. “Ethnocentrism, intercultural communication apprehension, intercultural willingness to communicate, and intentions to participate in an intercultural dialogue program: Testing a proposed model.”Communication Research Reports 20.1 (2003): 62-72.

[60] Dukes, Lauren Cairco, and Larry F. Hodges. ”Development of a scenario builder tool for scaffolded virtual patients.” Virtual Reality (VR), 2014 iEEE. IEEE, 2014.

[61] Gratch, Jonathan, et al. “User-state sensing for virtual health agents and telehealth applications.” MMVR. 2013.

[62] Hartholt, Arno, et al. “All together now.” Intelligent Virtual Agents. Springer Berlin Heidelberg, 2013.

[63] Kang, Sin-Hwa, and Jonathan Gratch. ”People like virtual counselors that highly-disclose about themselves.” The Annual Review of CyberTherapy and Telemedicine 167 (2011): 143-148.

[64] Rivera-Gutierrez, Diego, et al. “Getting the Point Across: Exploring the Effects of Dynamic Virtual Humans in an Interactive Museum Exhibit on User Perceptions.” Visualization and Computer Graphics, IEEE Transactions on20.4 (2014): 636-643.

[65] Gratch, Jonathan, Stacy Marsella, and Peter Carnevale. Modeling and Evaluating Emotions Impact on Cognition. UNIVERSITY OF SOUTHERN CALIFORNIA LOS ANGELES INST FOR CREATIVE TECHNOLOGIES, 2013.

[66] Rossen, Brent, and Benjamin Lok. “A crowdsourcing method to develop virtual human conversational agents.” International Journal of Human-Computer Studies 70.4 (2012): 301-319.

[67] Morbini, Fabrizio, et al. “FLoReS: A forward looking, reward seeking, dialogue manager.” Natural Interaction with Robots, Knowbots and Smartphones. Springer New York, 2014. 313-325.

[68] Thiebaux, Marcus, et al. “Smartbody: Behavior realization for embodied conversational agents.” Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[69] Boulic, R., Magnenat-Thalmann, N., Thalmann, D.: A global human walking model with real-time kinematic personification. The Visual Computer 6(6) (1990).

[70] Park, S.I., Shin, H.J., Shin, S.Y.: On-line locomotion generation based on motion blending. In: Proc. of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA), New York, NY, USA, ACM Press (2002) 105111

[71] Kallmann, M.: Scalable solutions for interactive virtual humans that can manipulate objects. In: Artificial Intelligence and Interactive Digital Entertainment (AIIDE), Marina del Rey, CA (June 1-3 2005).

[72] Liu, Y., Badler, N.I.: Real-time reach planning for animated characters using hardware acceleration. In Proc. of Computer Animation and Social Agents (CASA03). (2003) 8693.

[73] Kuffner, J.J., Latombe, J.C.: Interactive manipulation planning for animated characters. In: Proc. of Pacific Graphics00, Hong Kong (Oct 2000) poster paper.

[74] Baerlocher, P.: Inverse Kinematics Techniques for the Interactive Posture Control of Articulated Figures. PhD thesis, Swiss Federal Institute of Technology, EPFL (2001) Thesis number 2383.

[75] Tolani, D., Badler, N.: Real-time inverse kinematics of the human arm. Presence 5(4) (1996) 393401.

[76] Gu, E., Badler, N.: Visual attention and eye gaze during multipartite conversations with distractions. In: Proc. of Intelligent Virtual Agents (IVA06), Marina del Rey, CA (2006).

[77] E. Peacock and P. Wong, "The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal," Stress Medicine, vol. 6, pp. 227-236, 1990.

[78] C. Elliott, "The affective reasoner: A process model of emotions in a multi-agent system," Northwestern University Institute for the Learning Sciences, Northwestern, IL, Ph.D Dissertation 32, 1992.

[79] M. S. El Nasr, J. Yen, and T. Ioerger, "FLAME: Fuzzy Logic Adaptive Model of Emotions," Autonomous Agents and Multi-Agent Systems, vol. 3, pp. 219-257, 2000.

[80] Triola, M., et al. "A randomized trial of teaching clinical skills using virtual and live standardized patients." Journal of general internal medicine 21.5 (2006): 424-429.

[81] Kenny, Patrick, et al. "Virtual patients for clinical therapist skills training."Intelligent Virtual Agents. Springer Berlin Heidelberg, 2007.

[82] Conradi, Emily, et al. "Virtual patients in a virtual world: Training paramedic students for practice." Medical Teacher 31.8 (2009): 713-720.

[83] Wenger, Etienne. Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge. Morgan Kaufmann, 2014.

[84] Lazarus, Richard S. Emotion and adaptation. Oxford University Press, 1991.

[85] Lazarus, Richard S., and Susan Folkman. "Transactional theory and research on emotions and coping." European Journal of personality 1.3 (1987): 141-169.

[86] Begault, Durand, et al. "Applying spatial audio to human interfaces: 25 years of NASA experience." Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space. Audio Engineering Society, 2010.

[87] Litman, Diane J., et al. "Spoken versus typed human and computer dialogue tutoring." International Journal of Artificial Intelligence in Education 16.2 (2006): 145-170.

[88] Rizzo, Albert, et al. "SimCoach: an intelligent virtual human system for providing healthcare information and support." International Journal on Disability and Human Development 10.4 (2011): 277-281.

[89] Mehrabian, Albert. "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament." Current Psychology 14.4 (1996): 261-292.

[90] Lazarus, Richard S. "Progress on a cognitive-motivational-relational theory of emotion." American psychologist 46.8 (1991): 819.

[91] Scherer, Klaus R., Angela Schorr, and Tom Johnstone, eds. Appraisal processes in emotion: Theory, methods, research. Oxford University Press, 2001.

[92] Frijda, Nico H. "The laws of emotion." American psychologist 43.5 (1988): 349.

[93] Gebhard, Patrick. "ALMA: a layered model of affect." Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems. ACM, 2005.

[94] Brooke, John. "SUS-A quick and dirty usability scale." Usability evaluation in industry 189.194 (1996): 4-7.

[95] Shaikh, Mostafa Al Masum, Helmut Prendinger, and Ishizuka Mitsuru. "Assessing sentiment of text by semantic dependency and contextual valence analysis." Affective Computing and Intelligent Interaction. Springer Berlin Heidelberg, 2007. 191-202.

[96] Manning, Christopher D., Surdeanu, Mihai, Bauer, John, Finkel, Jenny, Bethard, Steven J., and McClosky, David. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55-60.

[97] Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining." LREC. Vol. 10. 2010.

[98] Turney, Peter D., and Michael L. Littman. "Measuring praise and criticism: Inference of semantic orientation from association." ACM Transactions on Information Systems (TOIS) 21.4 (2003): 315-346.

[99] Speer, Robert, and Catherine Havasi. "ConceptNet 5: A large semantic network for relational knowledge." The Peoples Web Meets NLP. Springer Berlin Heidelberg, 2013. 161-176.

[100] Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. Annual Review of Psychology 54: 547577.

[101] Kim, S. M., & Hovy, E. H. (2006). Identifying and analyzing judgment opinions. In Proceedings of HLT-NAACL, ACL (pp. 200207).

[102] Strapparava, C., Valitutti, A., and Stock, O. (2007). Dances with words. In Proceedings of IJCAI, Hyderabad, India. New York: Springer, pp. 17191724.

APPENDIX A: DR. CHESTR EVALUATION QUESTIONNAIRE

Consent Form

Project Title and Purpose:

You are invited to participate in a research study entitled "The Dr. CHESTR(Computerized Host Encouraging Students to Review) Game Show." The purpose of this study is to to study the effectiveness of having a virtual human with a distinct personality in such a way that it portrays a real game show host. We would also like to know if the participants can find and relate to the virtual human that we have created.

Investigators:

Raghavi Sakpal and Dale-Marie Wilson, Ph.D.

Office: Woodward 423B

Email: DaleMarie.Wilson@uncc.edu

Please contact Dale-Marie Wilson with questions regarding this research.

Description:You will be able to interact with the tutorial system by creating your own username and password. You will only be identified by your username. Information about your interaction with the tutorial system will be logged by the computer. You will be asked to fill out questionnaires before and after the session. The session will last about 30 minutes. Approximately 200 people will take part in this study.

Length of Participation:

Your participation in this project will require one laboratory session lasting approximately 30 minutes.

Risks and Benefits of Participation:

There are no anticipated risks associated with participation in this study.

Volunteer Statement:

You are a volunteer. The decision to participate in this study is completely up to you. If you decide to be in the study, you may stop at any time. You will not be treated any differently if you decide not to participate or if you stop once you have started.

Confidentiality versus Anonymity:

The data collected by the investigator will not contain any identifying information or any link back to you or your participation in this study. The following steps will be taken to ensure this anonymity: The data collected will be kept anonymous and confidential by randomly assigning a participant number for each participant and only referring to the data by the given participant number. In addition, names of the participants will not be collected. Any data that is documented on paper will be stored and locked in a cabinet for one year with access only given to the primary and co-investigators listed on this form. Any electronic data will be stored on a single computer protected by a password.

Fair Treatment and Respect:

UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. Contact the University's Research Compliance Office (704-687-3309) if you have any questions about how you are treated as a study participant. If you have any questions about the project, please contact Dale-Marie Wilson.

Participant Consent:

I have read the information in this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age, and I agree to participate in this research project. I understand that I will receive a copy of this form after it has been signed by me and the principal investigator.

Participant Signature:_____

Date:_____

Pre-Experiment Questionnaire

Age:_____

Gender:__________

Student/Faculty/Staff:__________

If student, list major:__________

Race/Ethnicity:

☐ Caucasian

☐ Hispanic

☐ African American

☐ Native American

☐ Pacific Islander

☐ Other:__________

Highest Degree Obtained:

☐ High School

☐ B.S.

☐ B.A.

☐ M.S.

☐ M.A.

☐ Ph.D.

☐ Other:__________

Disabilities: ☐ Yes ☐ No

Is English your native or second language? ☐ Yes ☐ No

For approximately how many years have you been using a computer?__________

Have you used an online tutoring system before? ☐ Yes ☐ No

If yes, about how many times have you used that system?

☐ 0 - 4 times

☐ 5 - 8 times

☐ 9 - 12 times

☐ More than 12 times

Which online tutoring system have you used before?__________

In the section below choose the response that most accurately describes you.

1. I frequently read computer magazines or other sources of information that describe new computer technology.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

2. I know how to recover deleted or lost data on a computer or PC.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

3. I know what a LAN is.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

4. I know what an operating system is.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

5. I know how to install software on a personal computer.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

6. I know what a database is.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

7. I am computer literate.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

8. I am good with computers.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

Post-Experiment Questionnaire

Please respond by circling the reaction that best reflects your reaction to the tutorial/review interface:

Terrible	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Wonderful
Frustrating	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Satisfying
Dull	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Stimulating
Usable	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Not Usable
Boring	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Fun

Please respond by selecting the reaction that best reflects your impressions:

1. The interface was easy for me to use.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

2. It was easy to get started.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

3. It was easy finding the image.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

4. I knew what to say during a task.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

5. I have a good understanding of how Dr. Chestr works.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

6. If you had errors, it was hard to recover from them.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

7. I was able to successfully complete the task.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

8. I was intimidated by the interface used.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

9. The interface I used helped me to complete the task.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

10. I knew what to say to Dr. Chestr.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

11. Dr. Chestr was fast enough in response to my question.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

12. Dr. Chestr worked as I expected it to during the task.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

13. I had problems understanding Dr. Chestr.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

14. Dr. Chestr had problems understanding me.

   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

15. I liked the appearance of Dr. Chestr.

    ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

16. I would have preferred a female agent.

    ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

17. I was confident that Dr. Chestr would be able to help me.

    ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

18. I would have preferred that Dr. Chestr would be able to help me.

    ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

19. I would use Dr. Chestr again.

    ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

20. Dr. Chestr would be easy to use by people who dont know a lot about computers.

    ☐ Strongly Agree ☐ Agree s ☐ Neutral ☐ Disagree ☐ Strongly Disagree

21. Dr. Chestr gave me appropriate answers.

    ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

I would improve Dr. Chestr by:____

____

Additional comments or suggestions on the interface used:__________________

____

APPENDIX B: VSP WITH MILD TBI QUESTIONNAIRE

Consent Form

Project Title and Purpose:

You are invited to participate in a research study entitled "Evaluation of Virtual Standardized Patient with mild TBI". Our research objective is to develop Virtual Humans (VHs) with the ability to portray emotions and generate behaviors based on their history, education, personal experiences and cognitive state of mind. To this effect we propose an agent architecture, called Culturally Modified Agent Architecture (CMAA), with the ability to generate autonomous VHs, whose behavior are driven by three main factors: Mood, Personality and Personal History. A VH prototype portraying a veteran with mild TBI (Traumatic Brain Injury), was developed as a screening tool for evaluation purposes. Our aim is to test the validity and believability of the VH prototype as veteran portraying symptoms of mild TBI and test the effectiveness of using the VH as a training tool to practice diagnostic evaluation in a clinical setting.

Investigator:

Raghavi Sakpal, Ph.D student

Email: rsakpal@uncc.edu

Phone: (704) 497-6256

Responsible Faculty: Dr. Dale-Marie Wilson

Email: DaleMarie.Wilson@uncc.edu

Please contact Dr. Dale-Marie Wilson with questions regarding this research.

Description:

During the research study session you will interact with two different versions of Virtual Standardized Patient (VSP) using a web based interface. For each session you will be asking the VSP a set a Questions provided within the interface tool. For each Answer provided by the VSP you will be asked to identify the symptoms and behavior portrayed by the VSP. You will be asked to fill out a set of questionnaires before and after the session evaluating the interface and the behavior of the VSP. Your study session would last for approximately 60 minutes.

Length of Participation:

Your participation in this project will require one laboratory session lasting approximately 60 minutes.

Risks and Benefits of Participation:

There are no anticipated risks associated with participation in this study. During this study you will benefit from exposure to image search technology that is typically inaccessible to the general public.

Volunteer Statement:

You are a volunteer. The decision to participate in this study is completely up to you. If you decide to be in the study, you may stop at any time. You will not be treated any differently if you decide not to participate or if you stop once you have started.

Confidentiality:

The data collected by the investigator will not contain any identifying information or any link back to you or your participation in this study. The following steps will be taken to ensure that data collected remains non-identifiable: The data collected will be kept anonymous and confidential by randomly assigning a participant number for each participant and only referring to the data by the given participant number. In addition, names of the participants will not be collected. Any data that is documented on paper will be stored and locked in a cabinet for one year with access only given to the primary and co-investigators listed on this form. Any electronic data will be stored on a single computer protected by a password.

Fair Treatment and Respect:

UNC Charlotte wants to make sure that you are treated in a fair and respectful manner. If you have any questions about how you are treated as a study participant, please contact the Universitys Research Compliance Office:

Phone: (704) 687-1871

Email: uncc-irb@uncc.edu

If you have any questions about the project, please contact the responsible faculty

Dr. Dale-Marie Wilson: Phone: (704) 687-7988

Email: DaleMarie.Wilson@uncc.edu

Participant Consent: I have read the information in this consent form. I have had the chance to ask questions about this study, and those questions have been answered to my satisfaction. I am at least 18 years of age, and I agree to participate in this research project. I understand that I will receive a copy of this document after I have indicated my willingness to participate in the study.

Pre-Experiment Questionnaire

- Demographic Information:

  1. Gender:

     ☐ Female

     ☐ Male

  2. Age Group:

     ☐ Under 18 yrs

     ☐ 18 to 24 yrs

     ☐ 25 to 34 yrs

     ☐ 35 to 44 yrs

     ☐ 45 to 54 yrs

     ☐ 55 to 64 yrs

     ☐ Age 65 or older

  3. Ethnicity (or Race):

     ☐ White

     ☐ Hispanic or Latino

     ☐ Black or African American

- ☐ Native American or American Indian
- ☐ Native Hawaiian or Pacific Islander
- ☐ Asian
- ☐ Other (Please Specify):_____

4. Education (Highest degree received):
   - ☐ Less than High School
   - ☐ High School Graduate (includes equivalency)
   - ☐ Diploma
   - ☐ Associates Degree
   - ☐ Bachelors Degree
   - ☐ Ph.D.
   - ☐ Graduate or Professional Degree
5. Are you currently enrolled in the MS of Nursing Science program?
   - ☐ No
   - ☐ Yes

   If Yes, Indicate your Major:
   - ☐ Family Nurse Practitioner
   - ☐ Adult Gerontology Acute Care Nurse Practitioner
   - ☐ Other (Please Specify):__________

- Knowledge & Experience with Mental Health Diseases:

1. How many years of nursing experience do you have?__________

2. Do you have prior nursing experience in Mental Health?

☐ No

☐ Yes

3. Do you have prior nursing experience in Neurology?

☐ No

☐ Yes

4. Do you have prior nursing experience in Rehabilitation?

☐ No

☐ Yes

5. Do you have prior nursing experience in TBI (Traumatic Brain Injury):

☐ No

☐ Yes

If Yes, Indicate your experience level regarding TBI and its evaluation:

☐ Classroom Knowledge

☐ Cared for a TBI patient

☐ Practiced Clinician

☐ Other (Please Specify):__________

- Experience using Online Learning Tool:

1. Have you ever used any online interactive learning tool?

☐ No

☐ Yes

If yes, please specify the details of the tool:__________

Post-Experiment Questionnaire Condition I

UNCCHARLOTTE

- Virtual Patient Evaluation

		Strongly Agree				Strongly Disagree
1	The interaction with the VSP was intuitive and enjoyable.	5	4	3	2	1
2	I was able to successfully complete the task/entire interaction with the VSP.	5	4	3	2	1
3	The VSP was fast enough in response to my questions.	5	4	3	2	1
4	The VSPs response were clear and easy to understand.	5	4	3	2	1
5	The VSP had difficulty understanding my questions.	5	4	3	2	1
6	The VSPs appearance was realistic.	5	4	3	2	1
7	The VSPs behavior was realistic.	5	4	3	2	1
8	The VSPs non-verbal behavior (facial expressions and gestures) were realistic.	5	4	3	2	1
9	The VSP was engaging in his conversation with you.	5	4	3	2	1
10	The VSPs symptoms were realistic and identifiable.	5	4	3	2	1
11	The VSPs symptom matched its response.	5	4	3	2	1
12	The VSPs emotions were realistic and identifiable.	5	4	3	2	1
13	The VSPs emotions matched its response.	5	4	3	2	1

- Qualitative Questions:

  1. What did you like about the VSPs behavior and appearance?

     ____

  2. Do you think the VSPs ability to portray emotions was more helpful?

     ____

  3. What changes would you like to see in the VSP?

     ____

  4. Any other helpful comments?

     ____

Post-Experiment Questionnaire Condition II

UNC CHARLOTTE

- Virtual Patient Evaluation

		Strongly Agree				**Strongly Disagree**
1	The interaction with the VSP was intuitive and enjoyable.	5	4	3	2	1
2	I was able to successfully complete the task/entire interaction with the VSP.	5	4	3	2	1
3	The VSP was fast enough in response to my questions.	5	4	3	2	1
4	The VSPs response were clear and easy to understand.	5	4	3	2	1
5	The VSP had difficulty understanding my questions.	5	4	3	2	1
6	The VSPs appearance was realistic.	5	4	3	2	1
7	The VSPs behavior was realistic.	5	4	3	2	1
8	The VSPs non-verbal behavior (facial expressions and gestures) were realistic.	5	4	3	2	1
9	The VSP was engaging in his conversation with you.	5	4	3	2	1
10	The VSPs symptoms were realistic and identifiable.	5	4	3	2	1
11	The VSPs symptom matched its response.	5	4	3	2	1
12	The VSPs emotions were realistic and identifiable.	5	4	3	2	1
13	The VSPs emotions matched its response.	5	4	3	2	1
14	The VSPs personality was realistic and identifiable.	5	4	3	2	1
15	The VSPs emotions & symptoms matched its personality.	5	4	3	2	1

- Qualitative Questions:

  1. What did you like about the VSPs behavior and appearance?

  2. Do you think the VSPs ability to portray emotions was more helpful?

  3. Do you think the VSPs ability to portray personality was more helpful?

  4. What changes would you like to see in the VSP?

  5. Any other helpful comments?

Post-Experiment Questionnaire

UNCCHARLOTTE

- Overall Review of the System:

Terrible	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Wonderful
Frustrating	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Satisfying
Dull	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Stimulating
Usable	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Not Usable
Boring	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Fun

- System Usability Scale (SUS):

		Strongly Agree				**Strongly Disagree**
1	I think that I would like to use this system frequently.	5	4	3	2	1
2	I found the system unnecessarily complex.	5	4	3	2	1
3	I thought the system was easy to use.	5	4	3	2	1
4	I think that I would need the support of a technical person to be able to use this system.	5	4	3	2	1
5	I found the various functions in this system were well integrated.	5	4	3	2	1
6	I thought that there was too much inconsistency in this system.	5	4	3	2	1
7	I would imagine that most people would learn to use this system very quickly.	5	4	3	2	1
8	I found this system very cumbersome to use.	5	4	3	2	1
9	I felt very confident using this system.	5	4	3	2	1
10	I needed to learn a lot of things before I could get going with this system.	5	4	3	2	1

- User Interface Evaluation:

1	Organization of information within the Interface.	Confusing 5	4	3	2	Very Clear 1
2	Reading characters on the screen..	Hard 5	4	3	2	Easy 1
3	Language of the Questions and Answers.	Confusing 5	4	3	2	Very Clear 1
4	Position of messages on the screen.	Inconsistent 5	4	3	2	Consistent 1
5	Position of buttons.	Inconsistent 5	4	3	2	Consistent 1
6	Learning actions associated with buttons.	Difficult 5	4	3	2	Easy 1
7	Ease to choose next set of action (question).	Difficult 5	4	3	2	Easy 1
8	Response time between question and answer.	Inconsistent 5	4	3	2	Consistent 1

- Virtual Patient Evaluation:

		Strongly Agree				**Strongly Disagree**
1	I observed differences in my interaction with the VSP in Condition 1 vs. Condition 2.	5	4	3	2	1
2	I observed differences in the VSPs behavior in Condition 1 vs. Condition 2.	5	4	3	2	1
3	I observed differences in the emotions portrayed by the VSP in Condition 1 vs. Condition 2.	5	4	3	2	1
4	I observed differences in the personality portrayed by the VSP in Condition 1 vs. Condition 2.	5	4	3	2	1
5	VSP portrayed in Condition 2 is better than VSP portrayed in Condition 1.	5	4	3	2	1
6	VSP with varying personality and emotions is more helpful in a diagnostic tool.	5	4	3	2	1
7	Using a VSP with varying behavior is more helpful than using an actor.	5	4	3	2	1

- Diagnostic Tool Evaluation:

		Strongly Agree				Strongly Disagree
1	I have a better understanding of patients suffering from mild TBI.	5	4	3	2	1
2	I have a better understanding of symptoms portrayed by mild TBI patients.	5	4	3	2	1
3	I have a better understanding of the behavior of mild TBI patients.	5	4	3	2	1
4	The VSP would be useful as a screening tool for mild TBI.	5	4	3	2	1
5	I would be comfortable interacting with a TBI patient after using the VSP interface.	5	4	3	2	1
6	The VSP would be useful as a good diagnostic tool for the nursing students.	5	4	3	2	1
7	I would use the VSP interface again.	5	4	3	2	1
8	Condition 2 would make a better diagnostic tool than Condition 1.	5	4	3	2	1

- Qualitative Questions:

1. What did you like about the VSP interface?

____

2. What changes would you like to see in the VSP interface?

____

3. Do you think a diagnostic tool with a VSP is better screening tool then interacting with an actor?

____

4. Do you think the VSP interface is a good diagnostic tool to practice mild TBI screening for nursing students?

5. Do you think you gained more knowledge about TBI? Would you be comfortable caring for & interacting with a TBI patient after using this tool?

6. Any other helpful comments?