

Stony Brook University



OFFICIAL COPY

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

© All Rights Reserved by Author.

Learning-based Intelligence for Socially Assistive Robots

A Thesis Presented

by

Junichi Terao

to

The Graduate School

in Partial fulfillment of the

Requirements

for the Degree of

Master of Science

in

Mechanical Engineering

Stony Brook University

May 2009

Stony Brook University
The Graduate School
Junichi Terao

We, the thesis Committee for the above candidate for the
Master of Science degree,
hereby recommend the acceptance of this thesis.

Prof. Goldie Nejat, Ph.D., Advisor
Department of Mechanical Engineering, Stony Brook University

Prof. Q. Jeffrey Ge, Ph.D., Chair
Department of Mechanical Engineering, Stony Brook University

Prof. Klaus Mueller, Ph.D., Member
Department of Computer Science, Stony Brook University

This thesis is accepted by the Graduate School.

Lawrence Martin
Dean of the Graduate School

Abstract of the Thesis

Learning-based Intelligence for Socially Assistive Robots

By

Junichi Terao

Master of Science

in

Mechanical Engineering

Stony Brook University

2009

Many countries in the world are now facing heightened social and economical demands in the health care environment due to the increase in the world's elderly population. The World Health Organization has reported that health care professionals such as doctors, nurses, therapists and other related professions are experiencing a significant shortage in their work force. The demands to fill job openings are substantially higher than the available supply of qualified people. In order to meet these health care demands,

advanced technologies such as socially assistive robots need to be implemented in health care environments. Socially assistive robots have the potential to assist health care staff by reducing their burden in performing repetitive and unskilled tasks. However, there are a number of research challenges that need to be addressed in order to develop such robots: (i) the robots need to have a high degree of autonomy, and cognitive and emotional capabilities, and (ii) the robots must be able to identify, understand and react to human intention and emotions. In order to meet these challenges, the objective of this research work is to develop intelligent controllers via the use of appropriate processing mechanisms for socially assistive robots. In particular, this research focuses on the unique development of a robotic emotional state module and a decision making module which work together to determine a robot's assistive behavior. The modules were integrated into the control architecture of a socially assistive robotic platform and experiments were conducted during one-on-one assistive interactions. The research work presented shows the potential utilization of these types of modules for robots engaged in human-robot assistive interaction scenarios.

TABLE OF CONTENTS

ABSTRACT	iii
TABLE OF CONTENTS	v
LIST OF FIGURES	viii
ACKNOWLEDGEMENTS.....	xii
CHAPTER 1 INTRODUCTION.....	1
1.1 Motivation.....	1
1.2 Literature Review.....	2
1.3 Research Problem Definition.....	9
1.4 Proposed Methodology and Research Tasks	10
CHAPTER 2 HRI CONTROL ARCHITECTURE	12
2.1 Related Control Architecture	12
2.2 A Robotic HRI Control Architecture.....	16

2.3 Chapter Summary	20
CHAPTER 3 ROBOT EMOTIONAL STATE MODULE	22
3.1 Related Emotional Models.....	24
3.2 Proposed Emotional State Module.....	25
3.3 Online updating.....	29
3.4 System reliability	31
3.5 Chapter Summary	32
CHAPTER 4 DELIVERATIVE LAYER	33
4.1 Reinforcement Learning	33
4.2 Q-Learning.....	35
4.3 A Proof of Concept Example.....	40
4.4 Chapter Summary	42
CHAPTER 5 EXPERIMENTS.....	43
5.1 Experimental Set-up.....	43
5.2 Implementation Scenarios	48
5.3 Chapter Summary	77

CHAPTER 6 CONCLUSIONS	78
6.1 Summary of Contributions.....	78
6.2 Discussion and Future Work.....	81
6.3 Final Concluding Statement.....	82
REFERENCES	83
APPENDIX A.....	94
APPENDIX B.....	97
APPENDIX C.....	104

LIST OF FIGURES

1-1 Rehabilitation Robotics Arm	5
1-2 Mobile Robot: SmartCane	5
1-3 Wheel Chair Robot: NavChair	5
1-4 Surgical Robot: Zeus	5
1-5 The Da Vinci Surgical System™	6
1-6 Medical Dispensing Robot	6
1-7 Paro	8
1-8 Ifbot	8
1-9 Kaspar	8
1-10 Pearl	9
2-1 Control Architecture	18
3-1 Emotional State Module	29
4-1 Reinforcement Learning Model	34
4-2 Learning-based Deliberative Layer	39

5-1 Social assistive robot, Brian	44
5-2 Examples of the behavioral actions of the robot during the interaction.....	48
5-3 Optimal Q -values for Experiment #2	52
5-4 Optimal Q -values for Experiment #3	53
5-5 Optimal Q -values for Experiment #4	53
5-6 Average number of iterations needed to satisfy all four drives	58
5-7 Emotional states of the robot during drive satisfaction	58
5-8 Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4	60
5-9 Transitions in the robot emotional state, the agent's accessibility level and the actions without updating method: (a) iterations from agent 1, (b) iterations from agent 2, (c) iterations from agent 3 and (d) iterations from agent 4	60
5-10 Average number of iterations needed to satisfy all four drives	63
5-11 Emotional states of the robot during drive satisfaction	63
5-12 Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4	65
5-13 Transitions in the robot emotional state, the agent's accessibility level and the actions without updating method: (a) iterations from agent 1, (b) iterations from agent 2, (c) iterations from agent 3 and (d) iterations from agent 4	65
5-14 Average number of iterations needed to satisfy all four drives	68

5-15 Emotional states of the robot during drive satisfaction	68
5-16 Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4	70
5-17 Transitions in the robot emotional state, the agent's accessibility level and the actions without updating method: (a) iterations from agent 1, (b) iterations from agent 2, (c) iterations from agent 3 and (d) iterations from agent 4	70
5-18 Average number of iterations needed to satisfy all four drives	73
5-19 Emotional states of the robot during drive satisfaction	74
5-20 Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4	75
5-21 Average number of iterations needed to satisfy all four drives	75
5-22 Emotional states of the robot during drive satisfaction	76
5-23 Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4	77
B-1 Average number of iterations needed to satisfy all four drives	99
B-2 Emotional states of the robot during drive satisfaction	99
B3 Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4	100
B-4 Average number of iterations needed to satisfy all four drives	101
B-5 Emotional states of the robot during drive satisfaction	101

B-6 Frequency of the accessibility levels of the agent when each drive was satisfied: (a)
Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4102

B-7 Average number of iterations needed to satisfy all four drives103

B-8 Emotional states of the robot during drive satisfaction103

B-9 Frequency of the accessibility levels of the agent when each drive was satisfied: (a)
Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4104

ACKNOWLEDGEMENTS

I would like to first thank my advisor Professor Goldie Nejat for her kind guidance for my research work and the thesis. I would also like to thank Dr. Maurizio Fico celli for his multiple discussions on my work. Also, thank you to Brian Allison, Nadia Gomez, Andrew Rosenfeld and Emmeline Kao for their help in developing the robotic platform. Lastly, I would like to thank you of my lab mates through out the 1 year and half of my study.

Chapter 1 Introduction

1.1 Motivation

Most countries in the world are now facing heightened social and economical demands in health care environments due to the increase in the world's elderly population. It is expected that the elderly will make up 20-32% of the population of a significant number of countries such as the U.S., Italy, Germany, and Japan in the next few decades [1, 2, 3]. These countries will need to invest billions of dollars to support and care for this age group in hospitals, nursing, veterans and private homes. What further escalates this issue is that the World Health Organization has also reported that health care professionals such as doctors, nurses, therapists and other related health professions are experiencing a significant shortage in their work force where the demands to fill job openings are substantially higher than the available supply of qualified people [4]. In order to combat these issues, advanced technologies should be implemented in patients' health care processes. For example, the use of assistive robots is expected to provide significant improvement in patients' safety and quality of care [5]. In particular, assistive robots have the potential to assist health care staff by reducing their burden in performing repetitive and unskilled tasks [6].

1.2 Literature Review

In general, robots designed for health care applications are placed into two categories: (i) non-interactive and (ii) interactive assistive robots [7]. The former group of robots gives aid or support to a human user without having any social interaction. Non-interactive robots include hands-on physical therapy and surgical robots as well as medication dispensing robots. The interactive robots, also known as the socially assistive group, are mainly designed to give aid or support through social interaction. The interactive robotics group includes robots that monitor patient health, help a patient with non-physical contact therapy, i.e., accompany patient on daily walking exercises, and provide companionship. In the following sections, the pertinent literature is reviewed for these two types of robotic aids.

1.2.1 Non-interactive robots

There are a number of non-interactive assistive robots that have been developed for health care applications including rehabilitation robots [8 -13], mobile aid and wheelchair robots [14-20], surgical robots [21 - 24] and medical dispensing robots [25 - 28]. A rehabilitation robotic arm, ARMin [9] is presented in Figure 1-1. ARMin is designed to assist in occupational therapy for patients with neurological and orthopedic injuries. ARMin has a semi-exoskeleton structure with position and force sensors. The robotic arm also has 6 degrees of freedom to allow it to assist in moving and positioning a person's arm during daily activities such as lifting, eating, and teeth brushing [9]. A mobile aid robot known as SmartCane, Figure 1-2, is designed as a walking aid for the elderly, in addition to being an active guide [14]. The SmartCane carries a CCD camera

to visually detect signposts placed strategically on the ceiling within the assistive living facility. The signposts that the CCD camera reads consist of three elements: Centerpiece Marker, Orientation Marker, and a unique pattern of Identification Markers. The Centerpiece Marker is used to identify the position of the signpost location, the Orientation Marker is used to orient the position of the user relative to the signpost location, and Identification Markers are used to recognize each signpost. As long as the CCD camera reads at least one signpost, it can detect the absolute position and orientation of the SmartCane. Obstacle avoidance is accomplished in a crowded environment using acoustic sensors. The SmartCane is also designed to approximate the user's directional intent by a 6 six-axis force-torque sensor, which measures the amount of force applied on the cane handle [14]. An example of a robotic wheelchair, NavChair, is presented in Figure 1-3 [17]. NavChair is designed to be used by people that have impairments in their motor, sensory, perceptual, and cognitive processes. NavChair has functions of obstacle avoidance, safe object approach, door passage, automatic wall following and keeping a straight path to minimize the use of motor and cognitive functioning to control the powered wheelchair [17]. Figure 1-4 presents the surgical robot, Zeus [21], which is designed for performing telesurgery remotely. It operates with the use of a master console and a patient console. Zeus is designed to allow surgeons to safely navigate a standard minimally invasive operation by using a digital sensory technique that filters out excess sensory input for smoother operation. The master console has a high-quality video screen to display the view of a surgical tool and patient from the endoscope. The patient console consists of three robotic arms and one instrument controller to manage the manipulation of graspers, scissors and other surgical instruments. In [22], a commercial robotic surgical

system the Da Vinci Surgical System™ is introduced, in Figure 1-5. The Da Vinci Surgical System™ is designed to operate surgeries through tiny incisions in patients' body, instead of fully exposing the operated part of the body in a traditional procedure. The Da Vinci Surgical System™ is the world's first robotic surgical system with high resolution vision (3 Dimensional Hi-Definition) to enhance visualization of surgical parts of patients for doctors. It has an accurate fingertip control called Endo Wrist, a motion scaling and tremor reduction system to enhance dexterity and precision of a doctor's surgical skills. The Da Vinci Surgical System™ has been utilized for a laparoscopic dismembered pyeloplasty surgery and performs complex procedures precisely [29]. In [25], a medical dispensing robot, SP 200, is introduced, Figure 1-6. This robot is designed to assist nurses and pharmacists to dispense prescriptions more efficiently and safely. By scanning the bar codes on a prescription receipt or vial, the robot determines the location of the medication and fills a vial held by its robotic arm with the exact number of prescribed pills. The pills are transported along a conveyer belt where a label is printed on the bottle containing the patient's name, instructions, and medication warnings. This robot helps reduce the wait time in acquiring medication from the pharmacy in a quick and efficient manner.

The assistive robots discussed above do not have any social abilities, and a large majority is contact-based robots or robots that perform crucial health tasks, there is an important concern in regards to the safety of the patients when integrating these robots in health care applications.

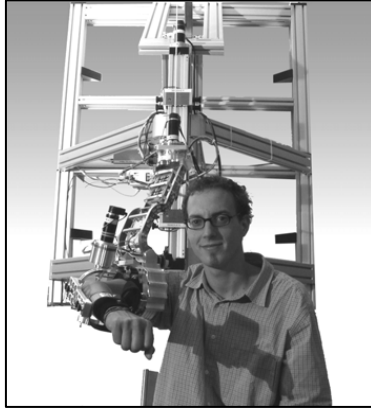


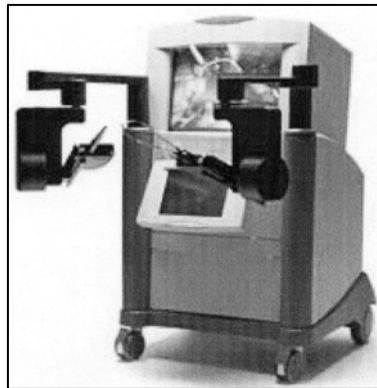
Fig. 1-1: Rehabilitation Robotics Arm [9].



Fig. 1-2: Mobile Robot: SmartCane [14].



Fig. 1-3: Wheelchair robot NavChair [17].



(a)



(b)

Fig. 1-4: Surgical robot: Zeus, (a) Master console and (b) Patient console [21].



Fig. 1-5: The Da Vinci Surgical System™ [22].

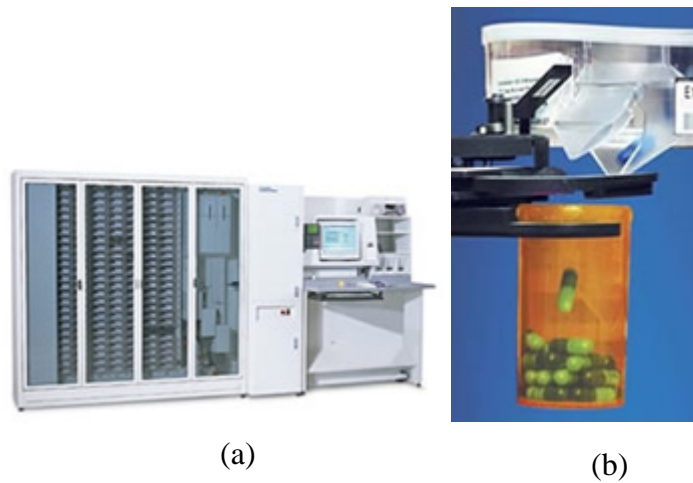


Fig. 1-6: Medical dispensing robot, SP 200: a) SP 200, and (b) robotic arm of SP 200 [25].

1.2.2 Interactive (Socially Assistive) Robots

A socially assistive robot can be defined as an interactive robot that develops *close and effective social interaction with a human user for the purpose of giving aid or support during rehabilitation, learning, and convalescence* [30].

In recent years, a number of socially assistive robots for health care applications in hospitals, medical and rehabilitation centers have been developed. For example, Paro, Figure 1-7, is a toy-like interactive robot modeled after a baby seal, having fur, whiskers, moving eyes and flippers. Paro has five types of sensors and responds to touch, sound, sight, and temperature changes [31]. Studies presented in [31] have shown that Paro has the potential of stabilizing patients' moods, reducing their dependency on the nursing staff and decreasing burden of the nursing staff. Similarly, a toy-like communication robot, Ifbot, Figure 1-8, has been developed [32]. Ifbot has 40 different facial expressions and several million patterns of word phrases to communicate with humans. It can be used as an aid to slow the progression of dementia in elderly people by providing continuous communication with the elderly via daily date, time and appointment schedule reminders. Ifbot also has a mental health diagnosis system, which asks specific questions and records the users' responses to detect symptomatic changes such as psychosomatic disease, neurosis, and depression. Kaspar, (Kinesics and Synchronization in Personal Assistant Robotics), Figure 1-9, is a child sized humanoid robot which is designed to improve basic social interaction skills in children with Autism using turn-taking and imitation games [33]. Pearl, Figure 1-10, is a socially assistive robot designed to be used to remind elderly people about their daily activities [34]. Pearl can express its emotions by using a cartoon-like face with moving eyes and eyebrows. Pearl also has a touch screen monitor for communication. Other socially assistive robots include CLARA [35], Patrol robot [36] and SIRA [37] that consist of a wheeled vehicle carrying a computer monitor projecting an image of a software agent or human.

The aforementioned socially assistive robots can provide significant improvements in hospitals, medical and rehabilitation centers through effective human-robot interaction. However, these robots were designed to give social aid or support without taking into account the emotions of the human they are interacting with. In order for socially assistive robots to accomplish effective task-oriented human-robot interaction (HRI), the robots need to understand and react to a human's emotion.



Fig. 1-7: Paro [31].



Fig. 1-8: Ifbot [32].



Fig. 1-9: Kaspar [33].



Fig. 1-10: Pearl [34].

1.3 Research Problem Definition

This research work focuses on addressing the design issues associated with the development of socially assistive robots for health care applications. In general, there are a number of challenges that need to be addressed in order to develop intelligent socially assistive robots for human-robot interaction scenarios. In particular, these robots need to have a high degree of autonomy, and cognitive and emotional capabilities, and be able to identify, understand and react to human intention and emotional state. It has been reported that the need for social intelligence in assistive robots is significantly important in a healthcare/eldercare environment [38]. However, the majority of socially assistive robots that have been developed today are unable to engage in emotion-based assistive interactions. Furthermore, the existing robots were designed with a focus on improving the robot's wellbeing. Namely, these robots do not consider the emotions of the people they are interacting with, hence, they are considered to be egocentric. In order for socially

assistive robots to accomplish effective HRI, the robots need to understand and react to human emotional states similar to human-human interactions. To overcome these limitations, the objective of this research work is to develop intelligent controllers via the use of appropriate processing mechanisms for the application of socially assistive robots. Namely, these processing mechanisms must be designed in a manner to allow these types of robots to learn from their environment and decide their appropriate assistive behavior during HRI.

1.4 Proposed Methodology and Research Tasks

The overall proposed methodology comprises the following components with corresponding reference to the Dissertation Chapters:

1. *HRI Control Architecture*

In Chapter 2, the generic HRI control architecture for socially assistive robots developed by our research group that is utilized in this thesis work is presented and discussed with respect to the current state-of-the art in HRI control architectures.

2. *Controller Design: Robot Emotional State Module*

In Chapter 3, the design of the robot emotional state module of the overall control architecture is presented. A Markov stochastic model is proposed as the main processing mechanism to effectively determine the robot's emotions during assistive HRI. On-line updating utilizing a positive influence factor is also discussed in

regards to on-line emotional processing. Furthermore, system reliability is utilized to allow quick changes for the robot emotional state.

3. *Controller Design: Decision Making Module*

In Chapter 4, the design of the decision making module of the overall control architecture known as the deliberative layer is presented. The deliberative layer is responsible for determining the task-driven behavior of a socially assistive robot that will best allow it to accomplish its given assistive tasks. For this module, the utilization and integration of a reinforcement learning method known as Q -learning is discussed.

4. *Implementation*

Chapter 5 presents the experiments performed to verify the effectiveness of the proposed robot control modules. In particular, the proposed processing mechanisms of the overall control architecture are integrated into a socially assistive robotic platform for one-on-one HRI.

Lastly, Chapter 6 presents concluding remarks on this research work, highlighting its main contributions and future work.

Chapter 2 HRI Control Architecture

2.1 Related Control Architectures

There have been a number of low-level control architectures designed for robots to mimic human emotional expressions [39-41], and high-level multi-module control strategies used to generate robotic emotions [32, 34, 36, 37, 42-52].

In [42], the CogAff (Cognition and Affection) control architecture was designed to model cognitive behavior. The architecture is divided into three layers: a deliberative layer, a reactive layer and a meta-management layer. The reactive layer produces a combination of state changes as a response to internal and external conditions. The deliberative layer is used to analyze, compare, evaluate and react to probable scenarios. Working as a reflective layer, the meta-management layer is responsible for self-observation or self-monitoring of internal states. The CogAff architecture was incorporated into a control architecture utilized to mimic human behavior for a service robot in [43]. The architecture consists of a number of modules: multi-modal interaction module, cognitive interaction module, environment intelligence module, and emotional interaction module. The multi-modal interaction module is responsible for perception of the environment, user identification and gesture recognition. The control architecture is designed to allow the robot to intelligently interact with humans through the use of the

cognitive interaction module, while the emotional interaction module consisting of the reactive and deliberative layers of the CogAff model is used to determine the robot's emotional behaviors. In regards to the robot's drive, the level of cognitive appraisal from an external stimulus influences the emotional expressions of the robot. This control architecture has yet to be implemented on a robotic platform to analyze its functionalities.

In [44], the Automatic-Deliberative (AD) architecture was proposed for the control of autonomous mobile robots. The AD architecture consists of two modules; the Automatic module and the Deliberative module. The Automatic module is responsible for low level control of actuators and sensors. The deliberative module is used for decision making capabilities and reasoning. Although the AD architecture presented in [44] is a human-based control architecture designed for autonomous robots, it does not take into account the use of emotions in its original design. In [45], an Emotional Control System (ECS) is integrated into the AD architecture providing an activity selection module used to determine goals and action tendencies based on the evaluation of a robot's wellbeing. The control architectures presented in both [44] and [45] has not yet been integrated into physical or simulated robotic platforms.

A computational model, Cathexis, was developed consisting of an emotion generation system and a behavior system to express emotional behavior within an object-oriented agent framework representing a child known as Simon [46]. External (environmental events) and internal (drives) are utilized as inputs into the model. Outputs from the model include motor system commands to represent the behavior of the agent. In particular, the behavior system coordinates with the emotional system to determine an appropriate behavioral output. Simon has three expressive components of emotion: (i)

prototypical facial expressions, (ii) body postures, and (iii) vocal expressions. In [47], an extended version of Cathexis for a pet robot is presented. This architecture consists of the following systems: perceptual, behavioral, drive, emotional and motor system. First, the perceptual system is responsible for visual and auditory processing of stimuli. The behavioral system acts as a distributed network of varying behaviors. The drive system urges the robot to take an action. The emotional system evaluates the stimuli for behavioral responses and future perceptions. Once the behavior has been chosen, appropriate motor actions are executed via the motor system.

In [48, 49], the behavioral control architecture of a robotic face, Kismet, is presented. This architecture consists of a number of systems: perceptual system, motivation system, attention system, behavior system and motor system. In the perceptual system, the sensory stimuli (i.e. visual images) are converted to meaningful information to guide the robot's behavior. The motivation system consists of two related sub-systems which implements drives and, emotions and expressive states. The drives act as internal representation of tasks, and the emotions and expressive states reflect the degree of task achievement. In this architecture, the drives act as a representation of the robot's agenda, and the emotions and expression states represent the satisfaction of the agenda achievement. The attention system determines saliency based on perception and motivation. In the behavior system, a chosen set of coherent actions are implemented. Finally, the motor system transfers behavior information to the actuators of the robot to display facial expressions. In [50], an emotional control architecture was developed for a humanoid robot. Within the control architecture, there are three main parts: behavior, emotion and cognition. In the behavior module, the behavior of the robot is determined

by sensor information or the emotional state of the robot. The emotion module enables emotional behavior based on drives and emotional expression. The cognitive module generates a plan to reach a certain goal by utilizing sensory information, emotional information and behavior information.

The assistive seal-like robot Paro has a behavior generation system consisting of proactive and reactive layers to generate the robot's proactive, reactive and physiological behaviors [51]. The proactive process, which is responsible for considering the internal state, stimuli, desire and rhythm of the robot, consists of two additional layers: behavior-planning layer and behavior generation layer. The behavior-planning layer determines the basic behavior pattern (i.e. poses, motions), and sends the behavior pattern to the behavior generation layer. The behavior generation layer generates control references for actuators to perform the determined behavior. The reactive layer is response to process sudden stimulation for immediate reactions of the robot. For the emotional interactive pet robot, KOBIE, an emotional expression system was developed [52]. The system consists of three components: Emotion Feature Generation Module (EFGM), Internal Status Generation Module (ISGM) and Behavior Decision Module (BDM). The EFGM generates factors that change the emotions of an emotional robot depending on the situation by considering relationship among emotional feature. The ISGM generates the internal status information using the emotion feature information. The BDM determines the behavior for the robot's emotion expression. The socially assistive robot, Pearl [34], has a hierarchical variant of a high level control architecture to determine the robot's behavior. For this architecture, a hierarchical partially observed Markov decision process (POMDP) was utilized to calculate optimal control actions under uncertainties. In this

architecture, the decision making is fully depended on probability distribution. The POMDP was also utilized in an assistive robot, SIRA, for learning, planning, and localization [37]

Even though a number of emotion-based behavior architectures have been developed in the literature, many of them have not been integrated into real-life scenarios and none have been designed for *task-driven* socially assistive robots. The aforementioned emotional behavior architectures have been designed to generate control commands based on the evaluation of the well-being of the robots. In fact, the affective state of the humans that the robots interact with is not taken into account in the control architecture. As mentioned before, the capabilities of understanding and reacting to human emotions is important especially in healthcare and elderly care environment. In order for socially assistive robots to accomplish effective HRI, the robots need to understand and react to human emotional states similar to human-human interactions. Therefore, it is useful to update control commands for socially assistive robots based on feedback from the emotional state of a human. This thesis focuses on developing appropriate processing mechanisms for a task-driven control architecture for socially assistive robots with an emphasis on utilizing human affective states to determine the robots' task-driven assistive behaviors.

2.2 A Robotic HRI Control Architecture

The main objective of a HRI control architecture is to incorporate artificial intelligence (AI) to enhance robot interactions with humans. The AI within the robot provides a function of autonomy, and recognition and identification of the user and

demonstrates human-like reactions and responses. These functions are especially important for socially assistive robots to understand human emotions and have real-time decision making capabilities to achieve effective interactions with humans. This section introduces the overview of the robot control architecture that has been developed by our research group, [53, 54, 55], and that will be utilized in this thesis work.

Figure 2-1 presents the overall HRI control architecture. The inputs into the architecture include the affective state of the person that the robot is interacting with (via the Human Mood Classifier module) and the robot's internal/external sensory information. The tasks that the robot needs to complete are stored in the long term memory module. The robot uses the human's ID to recognize the person. Once the robot identifies the person it is interacting with, tasks specific to that person will be sent to the drives module. The drives module will also consist of drives directly related to the robot's health (i.e., battery power, operation of motors) as updated from the robot's sensors. These drives will then be utilized to assist in determining the robot's emotional state via the robot emotional state module, and the output behavior via the reactive or deliberative layer. The emotional state is stored in the short term memory. In this architecture, the concept of deliberative and reactive layers, which was first developed in [45], was adapted for socially assistive robots. The deliberative layer determines a robot's behavior (i.e. actions, facial expression, and gestures) based on the human's affective states and its own emotional states. The reactive layer is utilized in determining a robot's behavior in case of interactive situations that require immediate response. The priority layer manages the final decision of the robot's behavior based on prioritizing the information during the interaction in regards to robot and human health and safety. Then

this information is relayed to the robot actuators to implement the appropriate behavior. This work focuses on non-contact socially assistive robots, the defined tasks that the robot needs to satisfy during interaction include providing reminders to patients, health and safety monitoring and providing companionship.

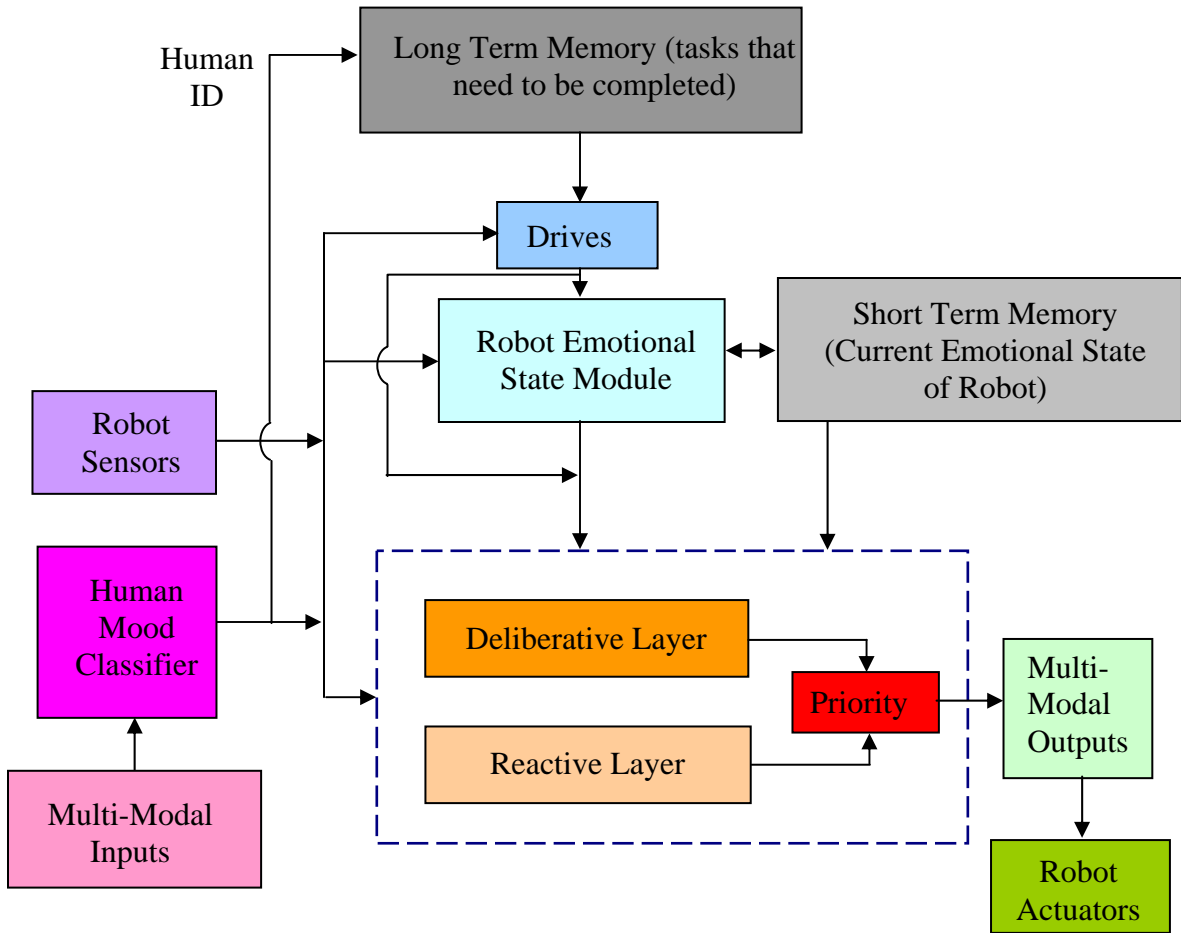


Fig. 2-1: Control Architecture [53, 54, 55].

To make HRI interactions more realistic, a robot's emotions need to be taken into account during the process of decision making for the following reasons: (1) the appropriate robot emotion has to correspond with the drive the robot must satisfy (i.e. the robot should not display the angry emotional state when providing companionship, and (2)

in the case when the robot is not able to satisfy a drive with a particular emotional state, the robot needs to adjust its emotion in order to decide on the best emotion to satisfy the drive. For example, during an interaction, a person may refuse to perform one of the daily tasks given by the robot while the robot is in one particular emotional state, the robot must change its emotion according to the situation in order to complete its task.

Furthermore, the robot's behavior should reflect the task it needs to complete and its emotional state should assist in the robot completing the task, unless the robot is physically incapable (e.g. it does not have enough battery power). Hence, the objective is not to have the robot mimic human emotions, but to use emotions to assist in determining the behavior necessary for the robot to accomplish its tasks.

Particularly, in this thesis, the development of the robot emotional state module and deliberative layer is focused.

2.2.1 Human Mood Classifier

Human mood state

In human mood classifier, the human mood state, which is utilized in the robot emotional state module and deliberative module, are determined. In the robot emotional state module, human mood state is required to determine the robot emotional state and in the deliberative layer it is required to determine appropriate robot behavior during HRI.

In order to determine the human mood state, the Nonverbal Interaction and States Analysis (NISA) of the Davis Nonverbal States Scale (DNSS) [56] is utilized in this control architecture to determine the affective state of a human during assistive interactions with a robot [53, 54, 55]. The DNSS is a nonverbal coding system, which has

been developed by Davis and Hadiks, for the use of therapy application [56]. Davis and Hadiks have found that a significant relationship between a patient's body movement pattern and his/her accessibility level during patient-therapist interactions. NISA is developed to identify a person's accessibility level, which indicates how much the person is willing to have a one-on-one conversation, by categorizing the accessibility levels into four different levels: Level I (least accessibility) to Level IV (most accessibility). In this architecture, the accessibility level is used to reflect the person's mood state. For example, when the person is in accessibility level I, it can be assumed that he/she is not in a good mood to have a conversation with the robot. On the other hand, when the person is in accessibility level IV, he/she is happy to interact with the robot. These accessibility levels are categorized according to the following body part poses: (i) trunk lean and orientation; and (ii) arm symmetry, location and orientation. The human accessibility values determined by the scale during the interaction will be utilized by both the robot emotional state module and the deliberative layer in order to assist in determining the robot's behavior.

2.3 Chapter Summary

Although there have been a number of emotional behavior architectures proposed in the literature, few have been the subject of extensive implementation and analysis and none have been designed specifically for assistive robotic applications. The type of processing mechanisms to be utilized in each layer of the control architecture is usually left as the sole responsibility of the designer of the agent/robot. The objective of this overall research is to design a complete control architecture that can be utilized effectively for task-driven socially assistive robots. In regards to this objective, the aim of

this thesis work is to devise and evaluate appropriate processing mechanisms for particular modules within this control architecture as described in Chapters 3 and 4.

Chapter 3 Robot Emotional State Module

Humans express emotions and communicate with each other mainly by using a combination of speech, body gestures and facial expressions [57]. Ekman has determined that there are 6 basic or universal emotions that humans can have: Happy, Sad, Angry, Afraid, Surprise and Disgust, [58]. These six emotions are universally recognized and expressed during human-human interaction.

In recent studies, many researchers have been developing the emotion in computer animated agents or robots to achieve emotional human-robot interactions. It has been reported that in order to accomplish such human-robot interaction, it is essential for the robot to have an ability to express its emotions [59, 60, 61]. In addition to the ability of emotional expression, the robots also need to have a capability to understand each other's emotion through speech, body gesture, and facial expression during HRI. A number of studies have been conducted in regard to emotion in robots, [49, 50, 52, 62-64], these main objectives to allow the robots to have emotion is performing natural interaction/communication with human/environment which includes enhancing general robot behaviors.

A 3-directional robotic head, Kismet has been developed by Cynthia Breazeal and colleagues at Massachusetts Institute of Technology (MIT), [48, 49]. Kismet has the 6 universal emotions (happy, sad, angry, fear, surprise, disgust) and simulates the emotions

through various facial expressions by movable eye, eyelid, lips and ears to participate in human social interaction.

In [62, 63], a humanoid robot, WE-4RII, was built as a platform to test their new mechanism which is able to interact naturally with human by using its 6 emotions - happy, neutral, sad, angry, disgust, and fear. WE-4RII expresses its emotion through facial expression and body gesture. WE-4RII was designed to communicate with human in human-like manner especially in medical or nursing care service with elderly people. KOBIE (Koala Robot with Intelligent and Emotion) is a koala-like emotional robot that was designed to give emotional support to patients through affective interaction by using its seven emotions (fear, surprise, joy, anger, sad, shame, and neutral). KOBIE expresses its emotion through facial expression, body gesture, and sounds, which are all associated with behavior of the robot [52]. KOBIE uses 20 force-sensing registers and 13 on/off sensors to react to actions of hitting, stroking, poking, tickling and embracing from a human. KOBIE was successfully implemented in a real life situation and the emotions KOBIE expressed were recognized by human users. In [64], an emotional based decision mechanism was developed. This mechanism was installed in a social robot to examine its effectiveness. The robot has explored a crowded area and interacted with humans using 3 emotions (happy, sad, and angry).

In order for the robots to achieve effective interactions with human, the robots must have an effective emotional processing mechanism to display its emotions. In this chapter, the proposed emotional processing mechanism, which emphasis on both human and robot emotional state to enhance HRI, are discussed.

3.1 Related Emotional Models

In recent researches, a number of conventional emotional models have been developed [65, 66]. These emotional models consist of a number of layers (i.e. perception, planning, and learning). Commonly, these types of emotional models focus on constructing a relation and rules between each layer. Hence, the robots need to follow these rules, and the robot's emotion will not have flexibility in dynamic or uncertain environments [67]. On the other hand, computational emotional strategies, which can construct human emotion models mathematically, have been developed in [68, 69, 70]. In particular, it has been proven that Markov modeling theories work well, especially at modeling dynamic environmental changes and uncertainties by adjusting state transition probabilities, [68, 69, 70]. In particular, such models with Markov modeling theories consist of the Markov property. A Markov Chain consists of set of state $S = \{s_1, s_2, \dots, s_r\}$ and transition probabilities. The process of the Markov chain starts in one of the states, and moves from one state to another. Each move from one state to another state is called a *step*. For example, if the Markov chain is in currently in state s_i at time step k and moves to another state s_j at time step $k+1$ with a probability called transition probability, denoted by p_{ij} [71]. There are a number of advantages to use Markov Chain: i) it does not require much computation at each time step, ii) if the probability distribution is known at time k , the distribution at the next time step ($k+1$) can be calculated and iii) the calculation can be accomplished by only multiplying a given probability distribution with a transition matrix [72].

In [70], a model to imitate human emotions was developed. This model consists of nodes and arcs. The nodes represent different emotional states and the arcs represent

the probability of moving out of each state. In [69], an emotional model for autonomous robots was developed. There are 4 emotional states (happy, fear, passivity, and anger) to express the robot's emotions. The model was designed to update its emotion at each time step based on current state and input state of the emotion. In [70], an emotional interactive model was developed using 6 basic human emotions (happy, sad, angry, fear, surprise, and disgust). These emotions are categorized into three groups: positive, negative, and neutral. The emotion is determined by stochastic matrices, which contain probabilities of each combination of the 6 emotions.

However, in [68, 69, 70], online updating of the Markov probabilities was not incorporated in these algorithms. In general, the model parameters of most Markov models are estimated using data with conditions that match as closely as possible to the expected experimental conditions. The performance of the Markov chain begins to degrade over time when a mismatch between the model parameters and the experimental conditions exist. This can happen during HRI when new humans interact with the robot. In order to minimize this mismatch, the Markov chain parameters must be adapted online to match new scenarios. Furthermore, although Markov Chain have been used to model humans emotion, they have not been integrated into robotic controllers for socially assistive robots.

3.2 Proposed Emotional State Module

In this work, a state space representation approach was utilized. This approach is able to express the robot's emotional state transition easily.

$$x_R(k+1) = Ax_R(k) + B_1x_H(k+1) + B_2d(k+1), \quad (3-1)$$

$$y_R(k) = f(x_R(k), \lambda), \quad (3-2)$$

where A , B_1 and B_2 represent state and control input. In particular, A matrix represents the robot emotional state transition probability distributions $\{a_{ij}\}$, B_1 represents the robot emotional state - human mood probability distributions $\{b_{ij}^1\}$ and B_2 represents the robot emotional state-drive probability distributions $\{b_{ij}^2\}$. $x_R(k)$, $x_H(k)$ and $d(k)$ represent state and control inputs defined as robot emotional state vector, human mood state vector and drive vector, respectively.

In this work, each of possible robot emotional state is determined by probability distributions, $x_R(k)$. $f(x_R(k), \lambda)$ represents a winner takes all threshold function utilized to determine the dominant emotional state, $y_R(k)$. As mentioned, the robot's emotional state, $x_R(k)$ is expressed by probability distribution, and the emotional state has following Markov property;

$$P(x_R(k+1) = e_i | x_R(k) = e_j) \text{ is the same for all } k \geq 1, \quad (3-3)$$

where $e = [e_1, \dots, e_m]$ is a vector of m robot emotional states and e_i is a single emotional state in vector e . m represents the possible robot's emotional state. In this thesis, 3 universal emotions (happy, sad, and angry) and an additional emotion neutral, are utilized to express the robot emotion, hence, $m=4$. Initial state probabilities are determined based on a uniform state distribution denoted by π . The transition matrices \hat{A} , \hat{B}_1 and \hat{B}_2 , which are transition counts moving from state i to state j , are defined. The transition matrices \hat{A} , \hat{B}_1 and \hat{B}_2 are associated with the transition probabilities matrices A , B_1 and B_2 respectively.

$$\hat{A} = \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12} & \dots & \hat{a}_{1n} \\ \hat{a}_{21} & \hat{a}_{22} & \dots & \hat{a}_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \hat{a}_{n1} & \hat{a}_{n2} & \dots & \hat{a}_{nn} \end{bmatrix}, \hat{B}^1 = \begin{bmatrix} \hat{b}_{11}^1 & \hat{b}_{12}^1 & \dots & \hat{b}_{1n}^1 \\ \hat{b}_{21}^1 & \hat{b}_{22}^1 & \dots & \hat{b}_{2n}^1 \\ \vdots & \vdots & \dots & \vdots \\ \hat{b}_{n1}^1 & \hat{b}_{n2}^1 & \dots & \hat{b}_{nn}^1 \end{bmatrix}, \hat{B}^2 = \begin{bmatrix} \hat{b}_{11}^2 & \hat{b}_{12}^2 & \dots & \hat{b}_{1n}^2 \\ \hat{b}_{21}^2 & \hat{b}_{22}^2 & \dots & \hat{b}_{2n}^2 \\ \vdots & \vdots & \dots & \vdots \\ \hat{b}_{n1}^2 & \hat{b}_{n2}^2 & \dots & \hat{b}_{nn}^2 \end{bmatrix}, \quad (3-4)$$

The emotional state transition probabilities matrices A , B_1 and B_2 can be simply calculated by taking the column proportions of the number of transition matrices \hat{A} , \hat{B}_1 and \hat{B}^2 . Hence, the transition probability matrices A , B_1 and B_2 are defined as following:

The emotional transition probabilities;

$$a_{ij} = P\left(x_R(k+1) = e_i \mid x_R(k) = e_j\right). \quad (3-5)$$

$$a_{ij} = \frac{a_{ij}^o}{\sum_{p=1}^n a_{ip}^o}. \quad (3-6)$$

The robot emotional state - human mood probabilities:

$$b_{ik}^1 = P\left(x_R(k+1) = e_i \mid x_H(k+1) = x_{Hp}\right), \quad (3-7)$$

$$b_{ik}^1 = \frac{b_{ik}^{1o}}{\sum_{q=1}^n b_{iq}^{1o}}, \quad (3-8)$$

Where $x_H = [x_{H_1}, \dots, x_{H_r}]$ is a vector of possible human emotions and x_{H_p} is a single human mood state in x_H .

The robot emotional state - drive probabilities:

$$b_{il}^2 = P(x_R(k+1) = e_i | d(k+1) = d_n), \quad (3-9)$$

$$b_{il}^{2o} = \frac{b_{il}^{2o}}{\sum_{r=1}^n b_{ir}^{2o}}, \quad (3-10)$$

Where $d = [d_1, \dots, d_q]$ is a vector of all possible robot drives and d_n is a single drive in vector d .

The probabilities a_{ij} , b_{ik}^1 and b_{il}^2 have the following properties:

$$0 \leq a_{ij}, b_{ik}^1, b_{il}^2 \leq 1, \quad (3-11)$$

$$\sum_{j=1}^n a_{ij}, \sum_{k=1}^n b_{ik}^1, \sum_{l=1}^n b_{il}^2 = 1. \quad (3-12)$$

An Overview of the proposed Markov Chain based emotional state model is presented in Figure 3-1. Once the robot's emotional state has been identified, it is utilized to assist in determining the robot's appropriate behavior.

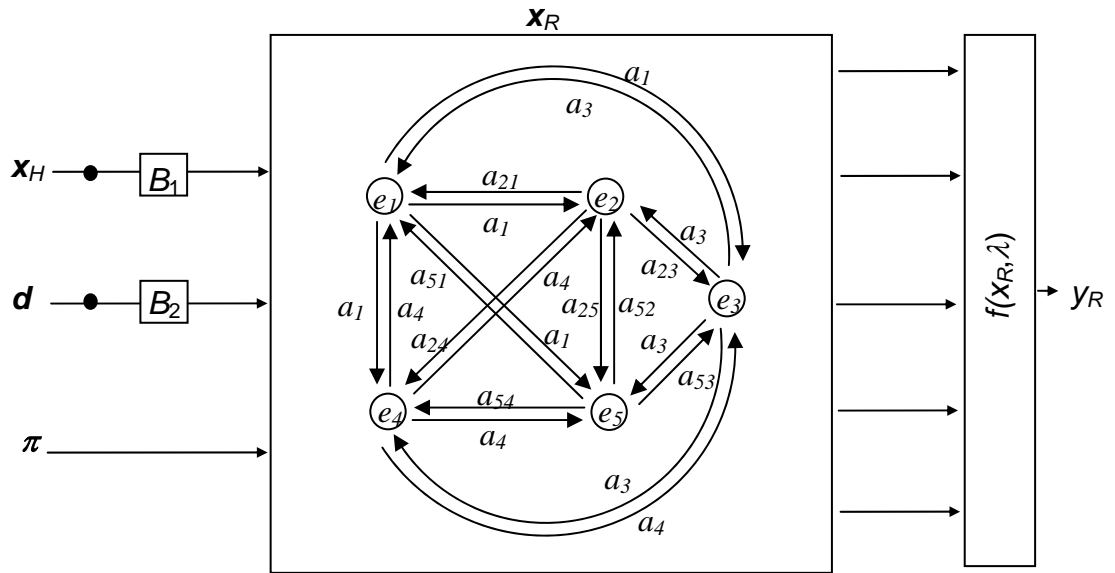


Fig. 3-1: Emotional State Module [53]

3.3 Online Updating

As previously mentioned, there might be some mismatch between the model parameters and the experimental conditions during HRI. This mismatch will affect the performance the Markov chain. Therefore, in order to minimize the mismatch, the parameters of the Markov chain must be update online.

In this work, online updating of the Markov chain transition probability matrices is achieved by using a positive influence factor, $\varepsilon_i(k)$. The positive influence factor is applied to the transition matrices to update for each time step. By using the positive influence factor, the robot will be able to display the most suitable emotional state to help satisfy the drives.

3.3.1 The Positive Influence Factor

The positive influence factor is distributed to the individual elements of the transition matrices, \hat{A} , \hat{B}^1 and \hat{B}^2 to update the transition probability matrices. There are two cases when the influence factor is applied to the transition matrices; i) when the robot successfully satisfies one of the drives, and ii) when the drives are not satisfied. For example, if the current robot emotional state is e_1 in time k and the robot stays in the same emotional state at time $k+1$ during which it satisfies one of the drives d_i while the human is in accessibility level defined by x_{Hi} , a positive influence factor would be applied to all corresponding elements of the transition matrices. On the other hand, if the drive is not satisfied, a positive influence factor is applied to the remaining elements of the transition matrices. The influence factor, herein, is defined for the two aforementioned cases:

1) When the drive is satisfied:

$$\varepsilon_i(k+1) = 1, \quad (3-13)$$

The positive factor is another transition 1 adding to the individual elements of \hat{A} , \hat{B}^1 and \hat{B}^2 .

2) When the drive is not satisfied:

$$\varepsilon_i(k+1) = \frac{1}{(m-1)}, \quad (3-14)$$

Where m is the total number of possible robot emotional states.

The positive influence factor can be treated as a type of “reward” for drive satisfaction or dissatisfaction. As can be seen from Equation (3-13) and (3-14), when the drive is satisfied the full transition reward 1 is given to a single element of matrices. On

the other hand, when the drive is not satisfied, the full transition reward needed to be distributed evenly to the remaining elements of matrices.

3.4 System Reliability

Utilizing the aforementioned approach to determining the emotional state can be quite effective. However, there is one situation in which the above approach may take a long time to determine an appropriate emotional state for the robot during interaction. This occurs when the probability of a dominant emotional state is significantly higher than the other emotional states, however, during interactions its use is unable to allow the robot to satisfy a drive. In particular, if the robot's dominant emotional state is ineffective in satisfying the drive at time k , then its probability of reoccurring at time $k+1$ within the same context (i.e., the same drive needs to be satisfied) should be reduced, hence providing opportunity for the other emotional states to be chosen. In order to address such situations, this work introduces the utilization of system reliability.

In general, system reliability is a simple Markov Model used to determine the reliability of a system based on the reliability of each individual module and the measured inter-modular transition probabilities [76]. It has been widely used in a number of software systems [73-76].

Herein, the reliability of emotion state i is denoted as $r_i(k)$ and represents the probability that this emotion state will satisfy the required drive d_i . The equation governing $r_i(k)$ is [75]:

$$r_i(k+1) = \frac{\hat{a}_{ij} - \Sigma_i(k)}{\hat{a}_{ij}}, \quad (3-15)$$

where \hat{a}_{ij} represents the transition counts moving from state i to state j and $\Sigma_i(k)$ is the number of times the robot's emotion state fails to transition. The system reliability is applied to the element of the transition probability matrices a_{ij} , and the reliability transition probability a_{ij}^r can be determined by;

$$a_{ij}^r = a_{ij} r_i(k+1) \quad (3-16)$$

The system reliability can be applied to a particular element of the transition matrix A only when the same emotional state becomes the dominant emotional state consecutively. As mentioned above, the system reliability is useful when the probability of a dominant emotional state is significantly higher than the other emotional states. Namely, the system reliability reduces the probability of the emotional state, which has the higher probability, so that other emotional states can have an opportunity to be the dominant emotional state. An example using system reliability is explained in Appendix A.

3.5 Chapter Summary

This chapter presents the emotional state module for a multi-layer control architecture for assistive robots. The module is based on a Markov state-space model. The online updating of the Markov probability matrices is accomplished through the use of a positive influence factor, $\varepsilon_i(k)$. System reliability, $r_i(k)$ was used to reduce the probability of a dominant emotional state of the robot, and give more opportunities to other emotional state to be a dominant emotional state.

Chapter 4 Deliberative Layer

4.1 Reinforcement Learning

Within the control architecture, the deliberative layer determines the task-driven behavior of the robot that will best allow it to accomplish its given tasks. To provide this kind of intelligence to the robot, this work focuses on the utilization and integration of reinforcement learning (RL) for the robot's decision making capabilities. Reinforcement learning is a real time, model free computational learning method in which an agent or robot learns how to accomplish its given tasks/goals by experiencing trial and error interactions with its environment [77,78]. Figure 4-1 presents a general reinforcement learning model. An agent/robot is in an environment, which consists of a set of states, S (s_0, s_1, s_2, \dots) and a set of actions, $A(\bar{a}_0, \bar{a}_1, \bar{a}_2, \dots)$. An agent/robot chooses an action \bar{a}_i in a state s_i to perform. After the action is performed, the agent/robot will move to a new state s_{i+1} . A reinforcement signal is received which indicates immediate reward r_i of this state-action interaction. Then the agent will choose another action \bar{a}_{i+1} to execute in state s_{i+1} . Reinforcement learning repeats this cycle of learning to choose actions that maximize the following [79];

$$r_0 + \gamma r_1 + \gamma^2 r_2 + \dots, \tag{4-1}$$

Where γ is known as a discount factor that denotes dependency of future interaction with the environment [79];

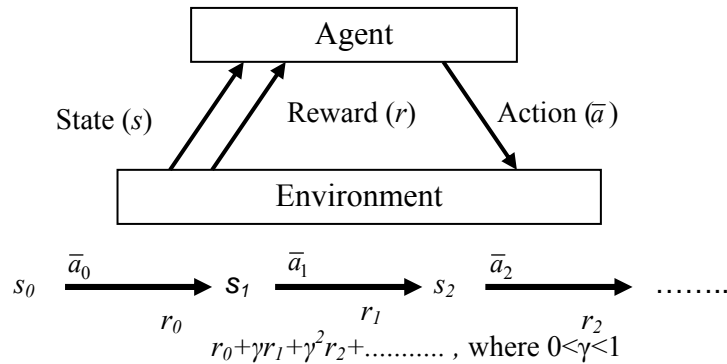


Fig 4-1; Reinforcement learning model [79]

Within reinforcement learning, there are a number of algorithms that can be applied; Temporal Difference (TD) learning which includes off-policy TD learning such as Q -learning and R -learning, and on-policy of learning known as State-Action-Reward-State-Action (SARSA). TD learning was first introduced by Samuel in 1959 and extended and modified by Sutton in 1988 [80]. TD learning has been proposed as a means of learning a wide variety of predictions about the interaction between an agent and its environment. However, conventional TD learning methods do not utilize trajectory data efficiently [80]. Hence, after an agent performs an update, the data of state transitions and rewards will be erased, [80]. Q -learning was introduced by Watkins and Dayan in 1992. Q -learning is considered as off-policy TD control algorithm, because the policy being learned might be different from the policy being executed. In Q -learning, an agent/robot learns from a state-action pair to a value called Q and compares the actions in the states and chooses the action, which has the highest Q value and executes it [79]. More detail about Q -learning is discussed in 4.2. R -learning is also a part of off-policy TD learning. As

mentioned above, Q -learning learns from the policy that maximizes the total rewards, on the other hand, R -learning learns from a policy that maximizes the average rewards, [81]. SARSA (state-action-reward-state-action) is an on-policy TD control algorithm. The SARSA learning algorithm is known as an on-policy algorithm because learning of the Q -value is fully depend on a policy it follows. SARSA learns an action-value function for a policy and estimates for the value of taking a state-action pair by sampling its future trajectory [82, 83].

4.2 Q-Learning

In this work, Q -learning was utilized to develop the deliberative layer for the robot's decision making ability. There are a number of advantages utilizing Q -learning: i) a priori information about the environment is not needed, and ii) the learning process is on-line [84]. Q -learning has been investigated in many simulated environments and real world scenarios [85, 86, 87]. In [85], Q -learning was utilized for obstacle avoidance for a mobile robot. The robot's action space is divided into six motion angles and for each motion angle, there are specific actions to take. After the motion angle is determined, Q -learning is used to choose the action with the highest Q -value stored in a rank table. After implementing an action, a reward (-1 to 0) is given to the robot to update the Q -values, which is determined based on how far the robot is from obstacles. In addition in [86], Q -learning is utilized for mobile robot navigation. The authors utilized a reward function, sparse, to update a Q -function. The sparse reward function 1 is assigned at goal and -1 at obstacles. By using this reward function, a distribution of Q -function on a lookup table will be obvious for the robot because the actions close to the goal would eventually

converge to higher Q -function than the ones around the obstacle. In [87], an intelligent transportation system (ITS), which provides the shortest path time to lead to a target place was developed by utilizing Q -learning and tested on Virtual Traffic Network (VTN). In this work, the most efficient path is found by using a Q -table, which contains the predicted time from current places to the target place. Hence, then it picks a path that has the smallest Q -values. The Q -values of each path are updated by getting a reward when the actual elapsed time is shorter than the predicted elapsed time.

As the previous studies show, Q -learning has potential applications in the field of agent/robot intelligent; however, it has mainly been applied to navigation problems and has not yet been applied and adapted to the field of socially assistive robots.

Q -learning can be applied to both deterministic and non-deterministic environments. For the deterministic environments, the next state is predictable, on the other hand, for the non-deterministic environments, the next state is not predictable.

Q-learning

In deterministic environments, the Q value is calculated as [79];

$$Q(s, \bar{a}) = r(s, \bar{a}) + \gamma \max_{\bar{a}'} Q(\delta(s, \bar{a}), \bar{a}'), \quad (4-2)$$

Where $r(s, \bar{a})$ is the immediate reward function for taking action \bar{a} from state s . γ is the discount factor and is set between 0 and 1 (γ represents independency of past experience) for future awards, i.e., a higher value places more emphasis on future awards).

$\delta(s, \bar{a})$ represents the state resulting from taking action \bar{a} to state s . The optimal action \bar{a} to take in state s is selected by following equation:

$$\pi^*(s) = \arg \max_{\bar{a}} Q(s, \bar{a}), \quad (4-3)$$

where $\pi^*(s)$ is called greedy policy and this policy selects an action that has the highest Q -value [79].

In order to estimate Q value for the first time the agent enters a new environment, $Q(s, a)$ is approximated by following rule:

$$\hat{Q}(s, a) \leftarrow r + \gamma \max_{\bar{a}'} \hat{Q}(s', \bar{a}'), \quad (4-4)$$

where $r = r(s, a)$ and s' is the state resulting from applying action \bar{a} to state s , and \bar{a}' are the actions applicable to the new state.

Q-learning in the Non-deterministic

On the other hand, for non-deterministic environments, the Q value can be expressed as [79];

$$Q_n(s, \bar{a}) = r(s, \bar{a}) + \gamma \sum_{s'} P(s'|s, \bar{a}) \max_{\bar{a}'} Q_n(s', \bar{a}'), \quad (4-5)$$

where $P(s'|s, \bar{a})$ is the probability of the resultant state based on the performed action. n represents the number of iterations. s' is the state resulting from applying action \bar{a} to state s , and \bar{a}' are the actions applicable to the new state.

In non-deterministic environment, if the rewards function is different at each time step, the same state-action pair will be selected. In order to avoid this situation, a training rule is needed to later the Q -values. The training rule is defined as [79];

$$\hat{Q}_n(s, \bar{a}) \leftarrow (1 - \alpha_n) \hat{Q}_{n-1}(s, \bar{a}) + \alpha_n [r(s, \bar{a}) + \gamma \max_{\bar{a}'} \hat{Q}_{n-1}(s', \bar{a}')], \quad (4-6)$$

where,

$$\alpha_n = \frac{1}{1 + \text{visit}_n(s, \bar{a})}, \quad (4-7)$$

The variable $\text{visit}_n(s, \bar{a})$ in Equation (4-7) represents the number of times action \bar{a} has been selected while the robot is in state s . Equation (4-7) is a learning rate, which decreases over time to allow for convergence.

Since human actions can be unpredictable when interacting with a robot, the non-deterministic Q -learning scheme is investigated for task-driven socially assistive robots in this work, where rewards are represented by probability distributions, Figure 4-2. Each state is defined by the mood y_{Hi} of the person, the robot's emotional state y_{Ri} and the drive d_i that needs to be satisfied. The socially assistive robot starts in a current state: i.e., $s_0(y_{Hi}, y_{Ri}, d_i)$, and will perform an action that will lead it to satisfy its dominant drive, d_i . For the Q -learning design, each state has multiple actions that can be implemented. In this work, Q -values of all the actions in a specific state are stored in a Q -value table. For example, if the robot is in state s_0 , the robot compares Q -values of all possible actions in the state and executes an action that has the maximum Q -value. Due to the uncertainty of the human-robot interaction, the drive may or may not be satisfied. The Q -value of the action will be updated and stored in the Q -value table after the action is performed.

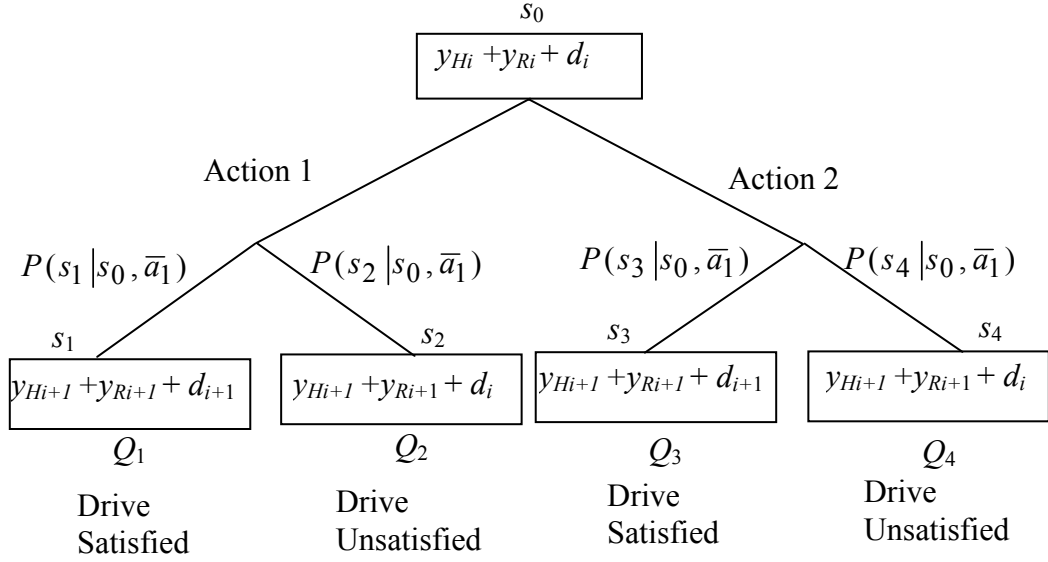


Fig. 4-2: Learning-based Deliberative Layer.

For example, assume Action 1, denoted by \bar{a}_1 in Figure 4-2, has a higher Q value and is implemented, the drive is satisfied, the robot will go on to state s_1 , ready to perform a new set of actions. State s_1 will consist of updated information regarding the robot's emotional state, the person's mood state and the next drive that needs to be satisfied (i.e., y_{Hi+1} , y_{Ri+1} , d_{i+1}). If the drive is not satisfied, the robot will move into state s_2 , where it will attempt to continue to satisfy its current drive, by updating its emotional state and the mood state of the human.

The immediate reward function in Equation (4-5) is defined by $r(s, \bar{a}) = w \times r_{di}$. Where w is a weight determined by the human mood state and r_{di} is a reward represented by the drive that needs to be satisfied. The value of r_{di} increases as the robot approaches its final drive. Particularly, in this work, r_{di} is defined as 50, 100, 150 and 200 as it approach to the final goal. The weight w is defined as 0.2, 0.4, 0.7 and 0.9 when the human mood state is level I, level II, level III and level IV respectively. As previously

mentioned, the drives will be directly related to tasks specific to the person the robot is interacting with and the robot's own health. The drives, in this work include, monitoring a human for health and safety, reminding the human of his/her daily activities, and providing companionship via verbal and non-verbal interaction in hospitals or nursing homes. In order for the robot to complete its tasks, the actions it must implement should be appropriate and fit within the context of the task at hand. The robot selects an action to perform based on the Q values in its current state.

4.3 A Proof of Concept Example

In this section, an example is given to illustrate the aforementioned methodology. The example is given by using actual values that are gained during experiments.

Mr. Brown has an appointment to see his doctor for his final check up today. 2 weeks ago he just underwent a surgery for stomach cancer and is to be discharged soon from the hospital. This is the final check up, which decides whether he can leave the hospital, or not. This is very important and it is imperative that he must see his doctor today. Since Mr. Brown is a few days away from discharging from the hospital, he has been in a good mood lately. In this situation, the human accessibility level y_{Hi} and current robot emotional y_{Ri} are determined as level III and Happy, respectively.

In this particular situation, the robot has four possible actions to take to remind him about the doctor's appointment:

Action 1: "Let's go to see your doctor" with the robot pointing in the direction of the doctor's room.

Action 2: "It's time to see your doctor now" with the root's arms to the side

Action 3: “You have a doctor’s appointment that you should go to” with the robot’s arm crossed”

Action 4: “Your doctor is waiting for you. You must go now to see him” with the robot’s body leaning forward.

Since human robot interaction is considered as non-deterministic Q -learning, the drive may or may not be satisfied. Therefore, there are two possible states that the robot can be after each action is implemented. Assume the robot is in current state of s_0 . Table 4-1 shows $Max Q$ -values and probabilities of each state.

Table 4-1: $Max Q$ -values and probabilities of each action.

	Action 1	Action 2	Action 3	Action 4
Max Q	338.7	280.5	215.6	285.6
Drive satisfied	Q1	Q3	Q5	Q7
Max Q	267.2	236.6	222.1	184.8
Drive unsatisfied	Q2	Q4	Q6	Q8
Probability of drive satisfaction	0.25 $P(s_1 s_0, \bar{a}_1)$	0.3 $P(s_3 s_0, \bar{a}_2)$	0.48 $P(s_5 s_0, \bar{a}_3)$	0.36 $P(s_7 s_0, \bar{a}_4)$
Probability of drive unsatisfaction	0.75 $P(s_2 s_0, \bar{a}_1)$	0.7 $P(s_4 s_0, \bar{a}_2)$	0.52 $P(s_6 s_0, \bar{a}_3)$	0.64 $P(s_8 s_0, \bar{a}_4)$

In this example, the reward function, r_{di} is set to 150. The robot past experiences are to significantly influence future experiences, hence, the value of the discount factor, γ is set to 0.8. The weight w , 0.7 is used when corresponding the human accessibility level III is determined.

The robot calculates the Q values, presented in Table 4-2, and decides on the action, which results in the maximum Q -value, i.e. action 1. Hence, action 1 is considered to be the most desirable action for the robot to take in this situation.

Table 4-2: Calculated Q -values.

	Action 1	Action 2	Action 3	Action 4
Q-values	333.1	304.1	280.3	281.8

4.4 Chapter Summary

In this Chapter, the main decision making module for a socially assistive robot is discussed. In particular, a reinforcement learning based approach known as Q -learning is utilized to determine the robot's task-driven behavior. Since during HRI, a human's actions can be unpredictable, the non-deterministic Q -learning method is adapted herein.

Chapter 5 Experiments

Several preliminary experiments were conducted to verify the proposed modules of HRI control architecture. Herein, the social assistive robot, Brian developed in [88, 89, 90] was utilized to simulate real human-robot interaction scenario. This chapter outlines the experimental set-up, procedure and results of these experiments. Furthermore, results for additional experiments can be found in Appendix B.

5.1 Experimental Set-up

5.1.1 Socially Assistive Robot

In this thesis, a socially assistive robot, Brian, that was developed in [88, 89, 90], was used to simulate real human-robot interaction scenarios. Brian is designed to have a number of functions to behave and act similar to a human from its waist up. The robot is able to communicate through: (i) a unique human-like face capable of displaying facial expressions, (ii) a 3 degrees-of-freedom (DOF) neck capable of expressing head gestures, and (iii) an upper torso consisting of a 2 DOF waist and two 4 DOF arms designed to mimic human-like body language [90]. This robot is also capable of showing its emotion through a unique human-like face [87]. Brian is also able to verbally communicate with a human by utilizing commercial text-speech software. Figure 5-1 shows the social assistive robot, Brian.



Fig. 5-1: Social assistive robot, Brian [88, 89, 90]

5.1.2 Software

All algorithms of the proposed modules within HRI control architecture were coded in C++ using Microsoft Visual Studio 2005 and implemented on a Pentium IV, 2.99GHz and 504MB of RAM personal computer.

5.1.3 Human Mood State

As mentioned in Section 2.2.1, in order to determine the mood state of a potential human interacting with the robot, the Nonverbal Interaction and States Analysis (NISA) of the Davis Nonverbal States Scale (DNSS) [56] was utilized. The human mood state is categorized by 4 levels (level I to Level IV) to measure a human's accessibility level during interaction. These accessibility levels are categorized according to the following body part poses: (i) trunk lean and orientation; and (ii) arm symmetry, location

and orientation. The human accessibility values determined by the scale during the interaction will be utilized by both the robot emotional state module and the deliberative layer in order to assist in determining the robot's behavior.

5.1.4 Robot Emotional State

The socially assistive robot can express four emotional states: happy, neutral, sad and angry during interaction. Furthermore, the following guidelines were utilized to ensure effective and appropriate HRI would take place between a human and the robot: (i) the robot should not change its emotional state from one extreme emotion to another, i.e., from happy to angry, and (ii) when the robot is providing companionship, it should not do so in an angry state.

5.1.5 Drives/Actions

In order to make the experiments as close to real-world interactive scenarios as possible, four situations that may be seen in a health care environment were defined as the robot's given drives. For each drive, there are four possible actions that the robot can perform. These four drives and actions are as follows;

Drive 1: To request users to perform a walking exercise

The robot's actions in order to satisfy this drive could be stating the following:

Action 1: "The weather is nice. Let's take a walk outside." While the robot's arms raised to the ceiling.

Action 2: “Lets’ go for a walk outside.” While the robot points towards the door.

Action 3: “I don’t like walking by myself, will you join me?” While the robot’s arm in a walking pose.

Action 4: “Walking is crucial to your overall wellbeing. therefore, you must go for a walk now.” While the robot’s arms open.

Drive 2: To remind users to take their medications.

The robot’s actions in order to satisfy this drive could be stating the following:

Action 5: “It’s time to take your medication.” While the robot points at its mouth.

Action 6: “Would you like to take your medication.” While the robot’s arm is on its forehead.

Action 7: “This medication is very important for your health.” While the robot points at the patient.

Action 8: “You must take your medication now.” While the robot’s one hand is pointing at the table with a medication bottle on it.

Drive 3: To remind users to go to see their doctor.

The robot’s actions in order to satisfy this drive could be stating the following:

Action 9: “Let’s go to see your doctor.” While the robot is pointing in the direction of the doctor’s office.

Action 10: “It’s time to see your doctor now” While the robot’s arm to the side.

Action 11: “You have a doctor’s appointment that you should go to.” While the robot’s arms are crossed.

Action 12: “Your doctor is waiting for you. You must go now to see him.” While the robot’s body is leaning forward.

Drive 4: To invite users to a one-on-one companionship.

The robot’s actions in order to satisfy this drive could be stating the following:

Action 13: “I have many interesting things to discuss with you. Would you like to join me in a nice conversation.” While the robot’s elbows are flexed and the palms up.

Action 14: “Let’s spend some time together.” While the robot’s hand resting on its hip.

Action 15: “I thought you would like to have my company right now” While the robot’s head is tilted to one side.

Action 16: “I am currently by myself, would it be okay if I join you.” While the robot’s hand is on its chest.

Figure 5-2 shows examples of the behavioral actions of the robot during the interaction.

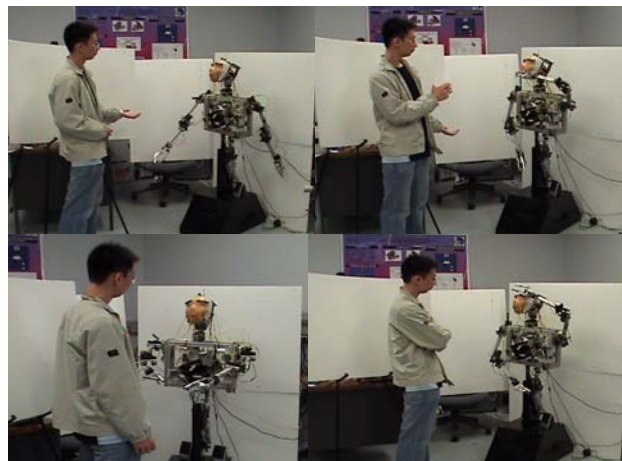


Fig. 5-2: Examples of the behavioral actions of the robot during the interaction.

5.2 Implementation Scenarios

Initial experiments were implemented to test the efficiency of the proposed processing modules in assistive HRI scenarios utilizing a simulated human agent.

This virtual agent was designed to mimic a human. The agent generates human-like answers based on its mood state (as defined by the accessibility levels) to respond to the robot's questions during interaction. The interaction between the robot and the virtual agent is generated automatically. The robot asks the virtual agent questions about the drives that need to be completed. Based on its accessibility level, the virtual agent responds. In these experiments, the accessibility level of the agent is determined from verbal content rather than nonverbal communication. The accessibility level of the agent is associated with the level of compliance. Particularly, the agent has four answers to choose from: "Yes", "Maybe", "I don't know", and "No." These answers reflect on the accessibility levels IV, III, II, and I, respectively (i.e. level IV is "Yes"). The accessibility level is weighted by a predetermined probability. If the agent is in an angry mood, the probability that the agent is in a low accessibility level is high. The overall probabilities of the agent being in any of the accessibility levels during this mood state are defined as:

$$\begin{aligned} P(x_H = x_{H1}) &= 0.4 \\ P(x_H = x_{H2}) &= 0.3 \\ P(x_H = x_{H3}) &= 0.2 \\ P(x_H = x_{H4}) &= 0.1 \end{aligned} \tag{5-1}$$

On the other hand, if the agent is in a happy mood, the probability that the agent is in the higher accessibility levels is also high:

$$\begin{aligned}
P(x_H = x_{H1}) &= 0.1 \\
P(x_H = x_{H2}) &= 0.2 \\
P(x_H = x_{H3}) &= 0.3 \\
P(x_H = x_{H4}) &= 0.4
\end{aligned} \tag{5-2}$$

Moreover, probability of the agent to be neutral mood can be expressed as:

$$\begin{aligned}
P(x_H = x_{H1}) &= 0.2 \\
P(x_H = x_{H2}) &= 0.2 \\
P(x_H = x_{H3}) &= 0.2 \\
P(x_H = x_{H4}) &= 0.4
\end{aligned} \tag{5-3}$$

The probability of the agent being in certain accessibility levels can be changed to reflect the mood state of the agent, therefore allowing different mood states to be designed for each agent.

Assume that the “degree of the mood” of the agent is defined by a series of integers:

$$x_H = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}, \tag{5-4}$$

The integers in Equation 5-4 representing the “degree of mood” are correlated with the four different accessibility levels via the following probability distribution determined for the agent angry mood state and happy mood state:

From Equation (5-1) (agent in an angry mood)

$$\begin{aligned}
P(x_H = x_{H1}) &= P(C_1) = 0.4 \\
P(x_H = x_{H2}) &= P(C_2) = 0.3 \\
P(x_H = x_{H3}) &= P(C_3) = 0.2 \\
P(x_H = x_{H4}) &= P(C_4) = 0.1
\end{aligned} \tag{5-5}$$

where $C_1=1, 2, 3, 4$, $C_2=5, 6, 7$, $C_3=8, 9$, and $C_4=10$.

From Equation (5-2) (agent in a happy mood)

$$\begin{aligned}
P(x_H = x_{H1}) &= P(C_5) = 0.1 \\
P(x_H = x_{H2}) &= P(C_6) = 0.2 \\
P(x_H = x_{H3}) &= P(C_7) = 0.3 \\
P(x_H = x_{H4}) &= P(C_8) = 0.4
\end{aligned} \tag{5-6}$$

where $C_5=1$, $C_6=2, 3$, $C_7=4, 5, 6$, and $C_8=7, 8, 9, 10$

From Equation (5-3) (agent in an neutral mood)

$$\begin{aligned}
P(x_H = x_{H1}) &= P(C_9) = 0.2 \\
P(x_H = x_{H2}) &= P(C_{10}) = 0.2 \\
P(x_H = x_{H3}) &= P(C_{11}) = 0.2 \\
P(x_H = x_{H4}) &= P(C_{12}) = 0.4
\end{aligned} \tag{5-7}$$

where $C_9=1, 2$, $C_{10}=3, 4$, $C_{11}=5, 6$, and $C_{12}=7, 8, 9, 10$.

During the experiments, the integer representation of the agent’s “degree of mood” was randomly picked to mimic an unknown HRI scenario. Based on the aforementioned relationship (i.e. Equations (5-4), (5-5), (5-6) and (5-7)) the accessibility level of the agent was then determined. A simple example of determining the accessibility level of agent is presented in Appendix C.

The interaction begins with the robot trying to satisfy drive 1, where it asks questions about completion of the drive. Once satisfied, the robot moves on to drive 2 and repeats this process until all four drives are satisfied. The robot does not move to the succeeding drive until the previous drive is satisfied. If drive 1 was not satisfied, the robot will continue attempting to satisfy the drive.

In this thesis, four types of experiments were implemented utilizing the virtual agent and one experiment using humans.

- Experiment #1 – Initialization of Markov Probabilities for the Robot Emotional State Module

In this experiment, the initial probabilities of Markov chains, which are required for the state-space matrices of the Markov-based robot emotional state module, is determined.

(Performed with both the virtual agent and humans)

- Experiment #2 – Implementation of the Emotional State Module and Deliberative Layer.

Experiment #2 was conducted to determine the behavior of overall control architecture without updating the Markov probability matrices.

- Experiment #3 - Implementation of Emotional State Module and Deliberative Layer with online updating of Markov probability matrices using the positive influence factor.

This Experiment was conducted to determine the behavior of overall control architecture with online updating of Markov probability matrices using the positive influence factor.

- Experiment #4 - Implementation of Emotional State Module and deliberative layer with the system reliability and online updating of Markov probability matrices using the positive influence factor.

This Experiment was conducted to determine the behavior of overall control architecture in the proposed assistive manner during HRI with the system reliability and online updating of Markov probability matrices using the positive influence factor.

- Experiment #5 Robot Design experiment with human subjects.

Experiment #5 was carried out to test how the robot's overall control architecture was effective in real HRI scenarios with humans. In particular, two types of experiments were conducted with the robot having a real human voice and a computer simulated voice.

Convergence of Q -value (Optimal Q -values)

Optimal Q -values for each action of the robot's drives were determined prior to implementing all the experiments. In this thesis, the virtual agent was utilized to interact with the robot to enhance the robot's experience. In order to determine these optimal Q -values, the training algorithm must visit every possible state a large number of times. The number of training iterations was set to be 10,000. All the Q -values obtained are stored in a look-up table. As the Q -values are determined from the robot-agent interactions over a number of iterations, the look-up table will eventually be filled with optimal Q -values (i.e., converged values).

Figures 5-3-5-5 present a sample of the optimal Q -values for Experiments #2, #3 and #4, respectively, after 10,000 iterations. Only 3 samples of the 16 optimal Q -values for each figure are shown since plotting all of the Q -values would make the figures difficult to read. All Q -values have converged. The three samples are for the robot implementing: Drive 1 under the emotional state of happy (Drive 1-Happy), Drive 3 under the emotional state of neutral (Drive 3-Neutral), and Drive 4 under the emotional state of happy (Drive 4-Happy). As can be seen in Figures 5-3-5-5, the Q -values converge to particular values after approximately 5000 iterations.

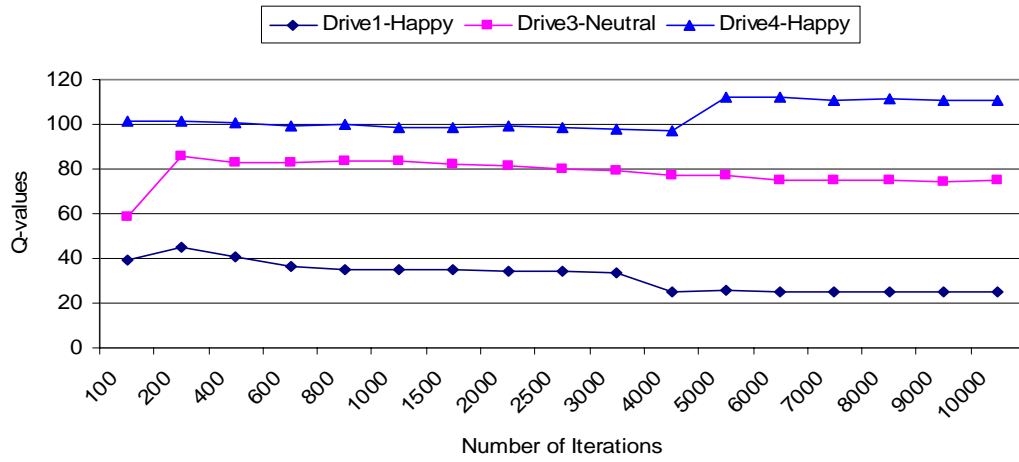


Fig. 5-3: Optimal Q -values for Experiment #2.

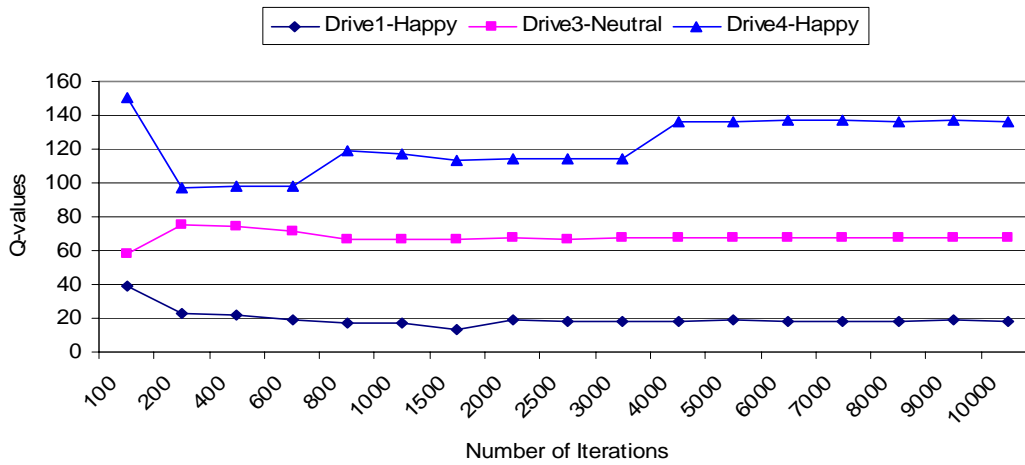


Fig. 5-4: Optimal Q -values for Experiment #3.

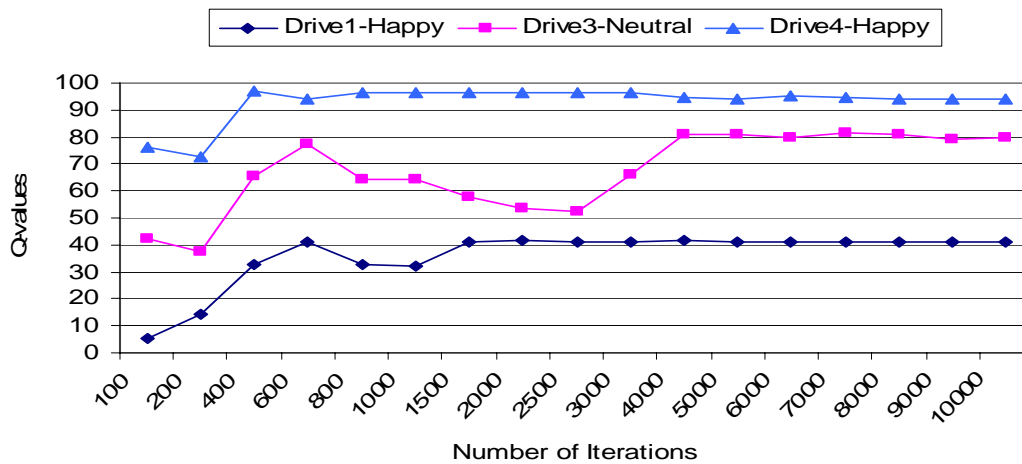


Figure 5-5: Optimal Q -values for Experiment #4.

5.2.1 Experiment #1: Initialization of Markov Probabilities for Robot Emotional State Module

Experimental Procedure

Experiment #1, the learning stage of the Markov chain, consists of using the Markov chain described in Section 3.2 to determine the initial probabilities required for the state-space matrices of the robot's emotional state module.

In this experiment, the initial probabilities of the Markov chain, i.e. matrix A , B_1 and B_2 , which is required for the state-space matrices of the robot's Markov-based emotional state module were determined. These matrices were simply constructed by assigning transition counts when the drives were satisfied. The matrices represent the state transition probabilities, the robot state – accessibility level probabilities and the robot state-drive probability, respectively. The rows of matrix A represent the current robot emotional state $e_i=[e_H, e_N, e_S, e_A]$ and the columns represent the robot emotional state $e_j=[e_H, e_N, e_S, e_A]$ for the next time step. Columns of matrix B_1 and B_2 represent the robot emotional state. The rows of matrix B_1 represent the accessibility levels: accessibility level I(x_{H1}) - accessibility level 4 (x_{H4}) and the rows B_2 represent drives: drive 1(d_1) - drive 4(d_4)

Experimental Results and Discussions

The results of the Initial learning stage of Markov probabilities are presented below,

Equation (5-8):

$$A = \begin{matrix} & e_H & e_N & e_S & e_A \\ \begin{matrix} e_H \\ e_N \\ e_S \\ e_A \end{matrix} & \begin{bmatrix} 0.356 & 0.269 & 0.250 & 0.238 \\ 0.267 & 0.385 & 0.150 & 0.285 \\ 0.266 & 0.192 & 0.400 & 0.048 \\ 0.111 & 0.154 & 0.200 & 0.429 \end{bmatrix} \end{matrix}$$

$$B_1 = \begin{matrix} & x_{H1} & x_{H2} & x_{H3} & x_{H4} \\ \begin{matrix} e_H \\ e_N \\ e_S \\ e_A \end{matrix} & \begin{bmatrix} 0.219 & 0.207 & 0.374 & 0.400 \\ 0.156 & 0.241 & 0.500 & 0.314 \\ 0.250 & 0.345 & 0.063 & 0.200 \\ 0.375 & 0.207 & 0.063 & 0.086 \end{bmatrix} \end{matrix}, \quad (5-8)$$

$$B_2 = \begin{matrix} & d_1 & d_2 & d_3 & d_4 \\ \begin{matrix} e_H \\ e_N \\ e_S \\ e_A \end{matrix} & \begin{bmatrix} 0.286 & 0.286 & 0.250 & 0.357 \\ 0.179 & 0.250 & 0.321 & 0.357 \\ 0.250 & 0.214 & 0.179 & 0.286 \\ 0.285 & 0.250 & 0.250 & 0.000 \end{bmatrix} \end{matrix}$$

As can be seen in matrix A, the highest probabilities are distributed diagonally. This shows that the robot has the general tendency to stay in its current emotional state. In matrix B_1 , when the accessibility level of a person is IV, the probability of the robot being in the emotional state of happy is the highest, because the robot tends to interact with humans well when they are more willing to interact with the robot. It can be seen in matrix B_2 , the probability of $B_2(4, 4)$ is zero, because in general, the robot should not be in an angry state when it is providing companionship.

5.2.2 Experiment #2: Implementation of Emotional State Module and Deliberative Layer

Experimental Procedure

This experiment was performed to determine the behavior of overall control architecture without updating the Markov probability matrices. First, the current robot emotional state, the agent's mood state and the drive that the robot needs to satisfy are inputted to the robot emotional state module to determine the robot emotional state for next time step. The robot emotional state is calculated by using the initial probabilities of the Markov chain, matrix A , B_1 and B_2 , which were obtained in Experiment #1. Once the robot emotional state for the next time step is determined, it is then sent to the deliberative layer to decide the most appropriate behavior of the robot.

Experimental Results and Discussions

Figures 5-6- 5-9 present experimental results for experiment #2. Figure 5-6 shows the average number of iterations needed to satisfy all four drives. Figure 5-7 presents the emotional states of the robot during drive satisfaction. Figure 5-8, shows the frequency of the accessibility levels of the agent when each drive was satisfied. Figure 5-9 depicts transitions in the robot emotional state, the agent's accessibility level and the actions that were performed during drive satisfactions. Another set of experimental results is shown

in Figure B-1-Figure B-3 of Appendix B.

In Figure 5-6, it took 2-3 iterations to satisfy the robot's required drives. It can be seen in Figure 5-7, all four emotional states were utilized to satisfy the drives. In total, the drives were satisfied 66 percent of the time when the robot was in the emotional state of happy, 32 percent of the time when the robot was in the emotional state of neutral, and 2 percent of the time when the robot was in the emotional state of sad, respectively. The emotional state of angry was not used to satisfy the drive because the probability of being angry is significantly lower than other 3 emotional states. In Figure 5-8, a greater number of drives were satisfied with the emotional state of happy and neutral when the agent was in accessibility level of IV or III. The robot was still able to satisfy the drives by using all four emotional states when the agent was in accessibility level of I or II. As mentioned in Section 5.1.5, each drive has 4 possible actions to take. Herein, *a1-a4*, *a5-a8*, *a9-a12* and *a13-a16* in Figure 5-9 represents different actions for drive 1, drive 2, drive 3, and drive 4, respectively. The roman numbers (I-IV) represent the accessibility levels of agent. In Figure 5-9(d), it took at most 13 iterations to satisfy all four drives. The average number of iterations that were required to satisfy all four drives was 11 iterations. The robot was in the emotional state of happy most of the time because the probability of the emotional state of happy was significantly higher than other emotional states, which can be seen in Equation 5-6. Even though the robot stayed in the emotional state of happy and neutral for a majority of the iterations, the robot was able to satisfy the drive by performing different actions.

Figure B-1 to Figure B-3 in Appendix B show similar results to Figure 5-6 to Figure 5-8. In Figure B-1, it took 2 or 3 iteration to satisfy the required drives. In Figure B-2, the robot utilized the emotional state of happy the most to satisfy all four drives. Similar to Figure 5-7, 71 percent, 27 percent and 1 percent of the drives were satisfied with the emotional state of happy, neutral and sad, respectively in Figure B-2. Same as Figure 5-8, Figure B-3 shows that the robot was able to satisfy the drives a significant amount of time with the emotional state of happy even when the agent was in accessibility level I.

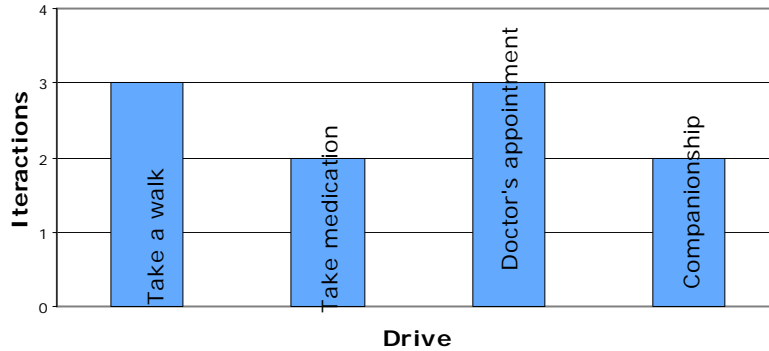


Figure 5-6: Average number of iterations needed to satisfy all four drives.

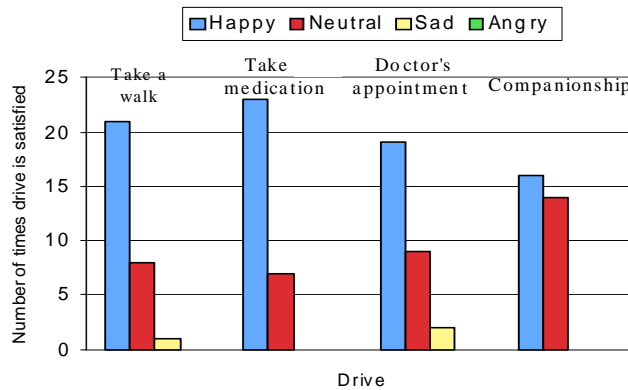
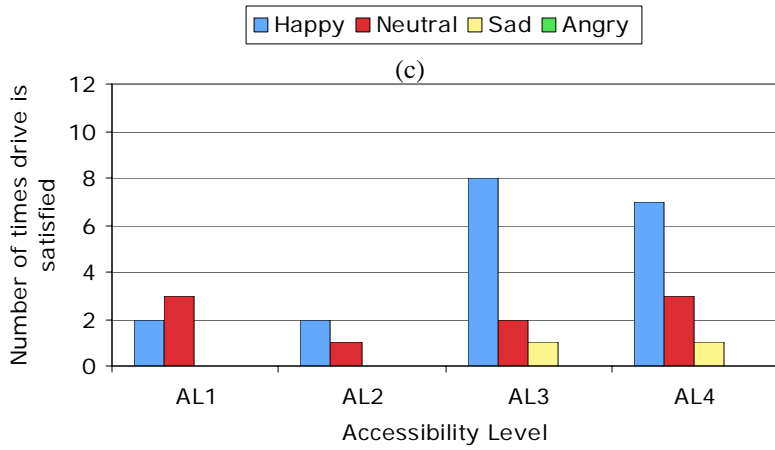
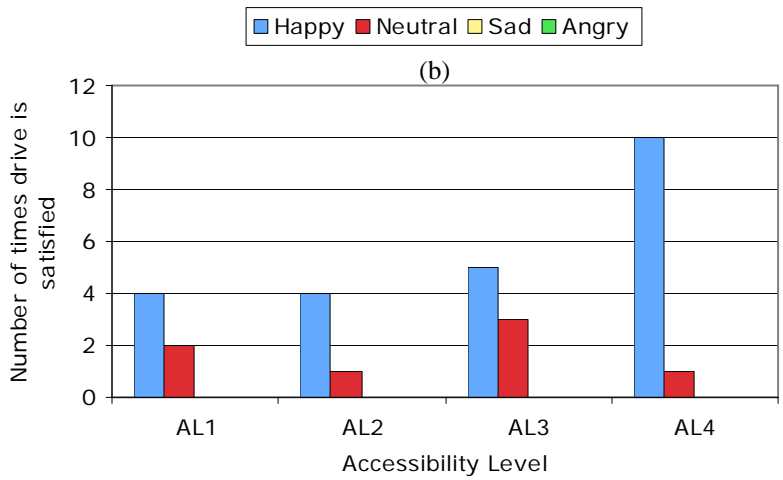
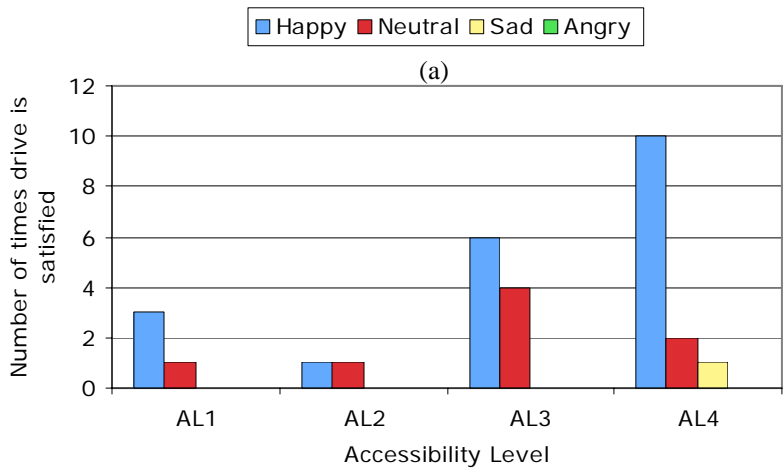


Figure 5-7: Emotional states of the robot during drive satisfaction.



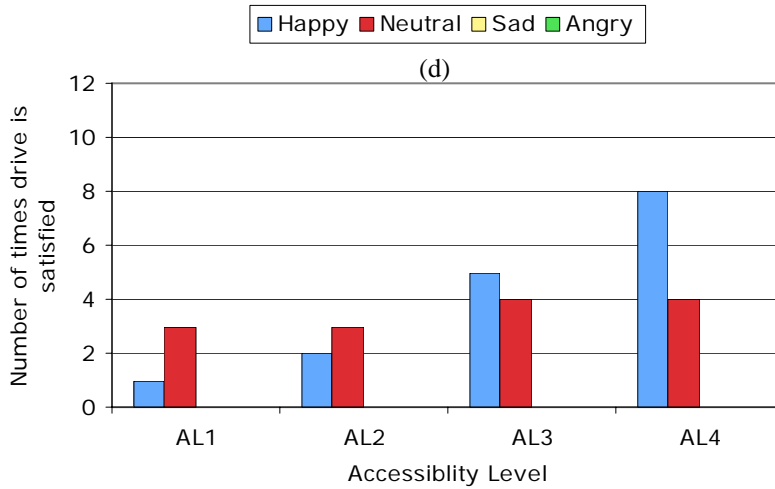


Figure 5-8: Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

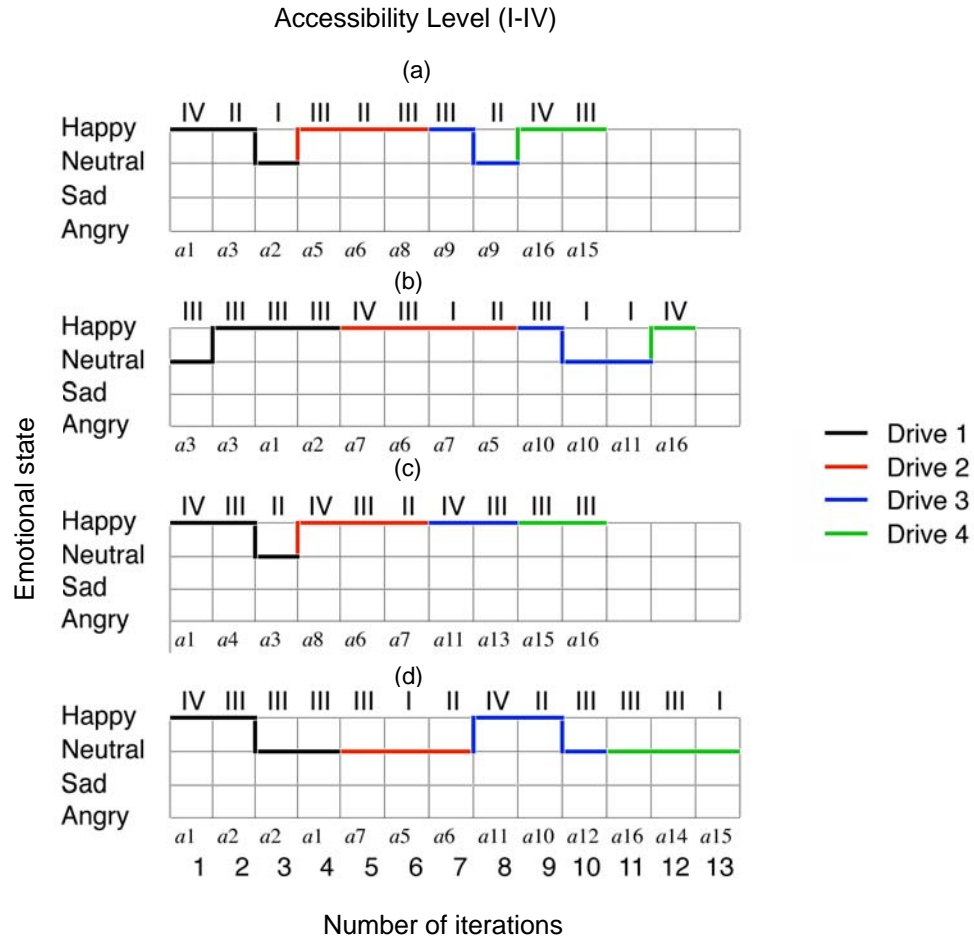


Fig 5-9: Transitions in the robot emotional state, the agent's accessibility level and the actions without updating method: (a) iterations from agent 1, (b) iterations from agent 2, (c) iterations from agent 3 and (d) iterations from agent 4.

5.2.3 Experiment #3: Implementation of Emotional State Module and Deliberative Layer with online updating of Markov probability matrices using the positive influence factor

Experimental Procedure

The implementation procedure of Experiment #3 is similar to Experiment #2. However, during this experiment, a positive influence factor, $\varepsilon_i(k)$, as described in Equations (3-13) and (3-14) was used to allow for online updating of the robot's emotional model. The influence factor was applied to corresponding matrix elements when the robot successfully satisfied one of its drives. On the other hand, if the drive was not satisfied, a positive influence factor was applied to the remaining transition elements described in Section 3.3.

Experimental Results and Discussions

Figures 5-10- 5-13 present experimental results for experiment #3. Figure 5-10 shows the average number of iterations needed to satisfy all four drives. Figure 5-11 presents the emotional states of the robot during drive satisfaction. Figure 5-12, shows the frequency of the accessibility levels of the agent when each drive was satisfied. Figure 5-13 depicts transitions in the robot emotional state, the agent's accessibility level and the actions that were performed during drive satisfaction. Another set of experimental results is shown in Figure B-4-Figure B-6 of Appendix B.

In Figure 5-10, it took 2-3 iterations to satisfy the robot's required drives. It can be seen in Figure 5-11, all four emotional states were utilized to satisfy the drives. In total, the drives were satisfied 58 percent of the time when the robot was in the emotional state of happy, 31 percent of the time when the robot was in the emotional state of neutral, 10 percent of the time when the robot was in the emotional state of sad and, 1 percent of the time when the robot was in the emotional state of angry, respectively. Similar to Experiment #2, in Figure 5-12, the majority of the drives were satisfied when the agent was in accessibility level of III and IV. In particular, the robot was able to satisfy the drive with the emotional state of angry when the agent was in accessibility level I in Figure 5-12(b) and accessibility level III in Figure 5-12(c). In Figure 5-13, actions $a1-a4$, $a5-a8$, $a9-a12$ and $a13-a16$ represent different actions for drive 1, drive 2, drive 3, and drive 4, respectively. The roman numbers (I-IV) represents accessibility level of agent. In Figure 5-13(d), it took at most 12 iterations to satisfy all four drives. The average number of iterations required to satisfy all four drives was 10 iterations. Figure 5-13 shows that the robot utilized all four emotional states (happy, neutral, sad, angry) to satisfy the drives. This experiment shows that the robot was able to utilize its four emotional states efficiently in addition to performing different actions. For example, in Figure 5-13(c), when drive 1 was implemented, the robot changed its emotional state from happy-happy-sad-sad according to the accessibility level transitioned from level III-III-II-I. The robot was able to satisfy the drive by performing action $a1$, $a3$, $a2$ and $a4$, respectively.

Similar experimental results can be seen in Figure B-4 to Figure B-6 in Appendix B.

Figure B-4 shows that, in total 2 to 3 iterations was required for the robot to satisfy the drives in Figure B-4. Figure B-5 shows that all four emotional states were utilized to satisfy the drives. Similar to Figure 5-11, 60 percent, 28 percent, 10 percent and 1 percent of drives are satisfied with the emotional state of happy, neutral sad and angry, respectively in Figure B-5. a grater number of the drives were satisfy when the agent was in accessibility level IV or III with the robot emotional state of happy and neutral in Figure B-6. Figure B-6 (a) and (b) show that the robot was able to satisfy Drives 1 and 2 with emotional state of angry.

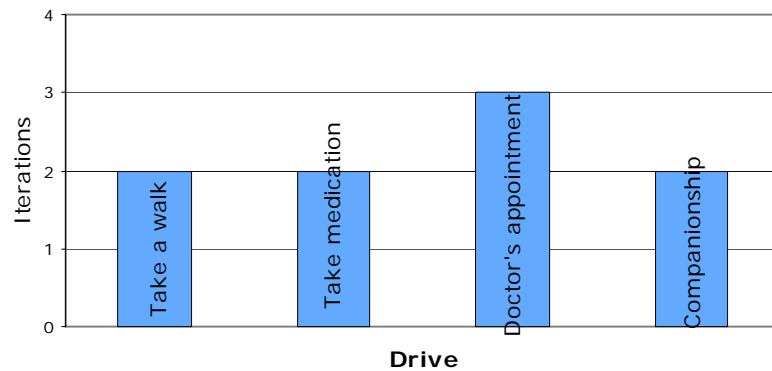


Figure 5-10: Average number of iterations needed to satisfy all four drives.

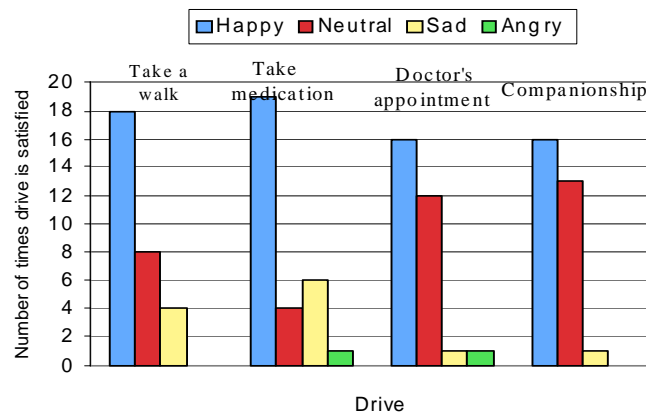
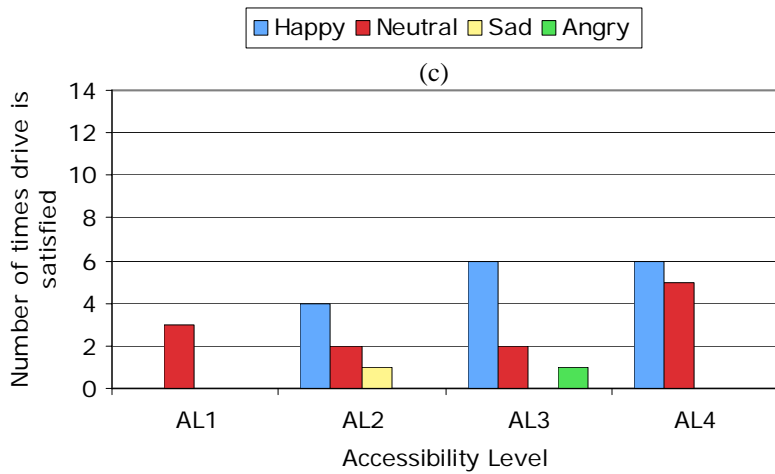
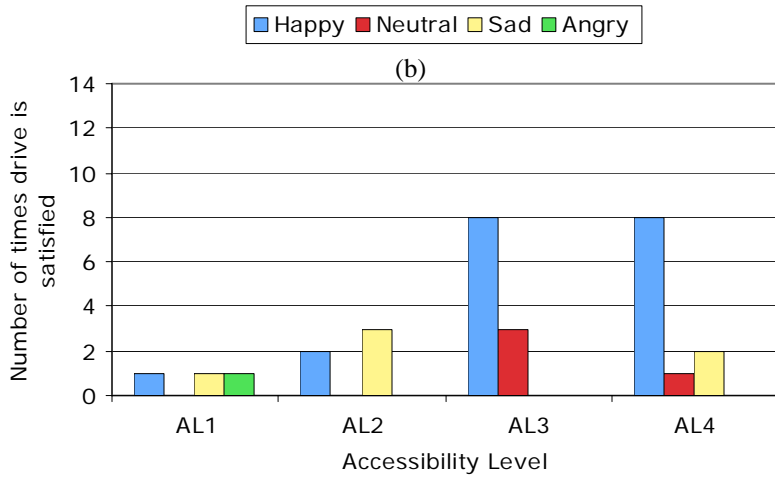
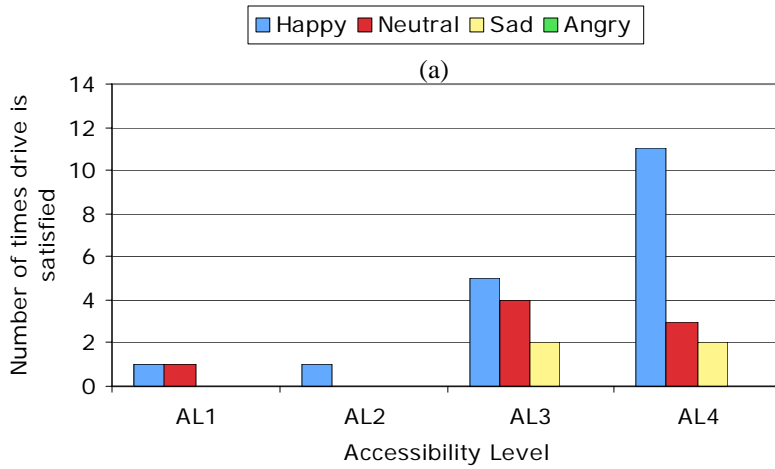


Figure 5-11: Emotional states of the robot during drive satisfaction.



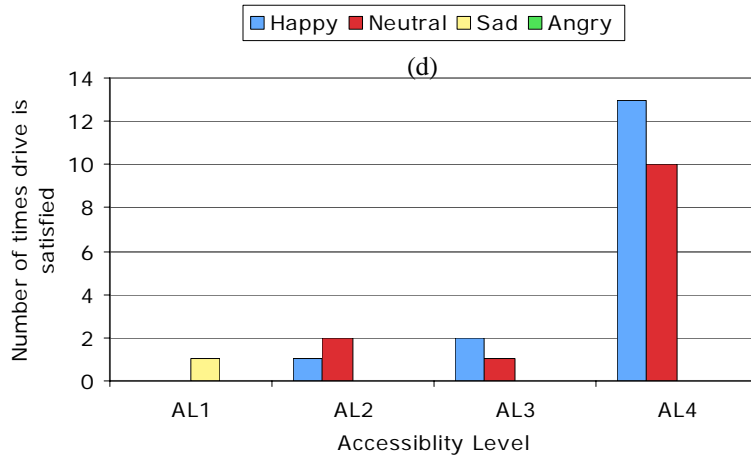


Figure 5-12: Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

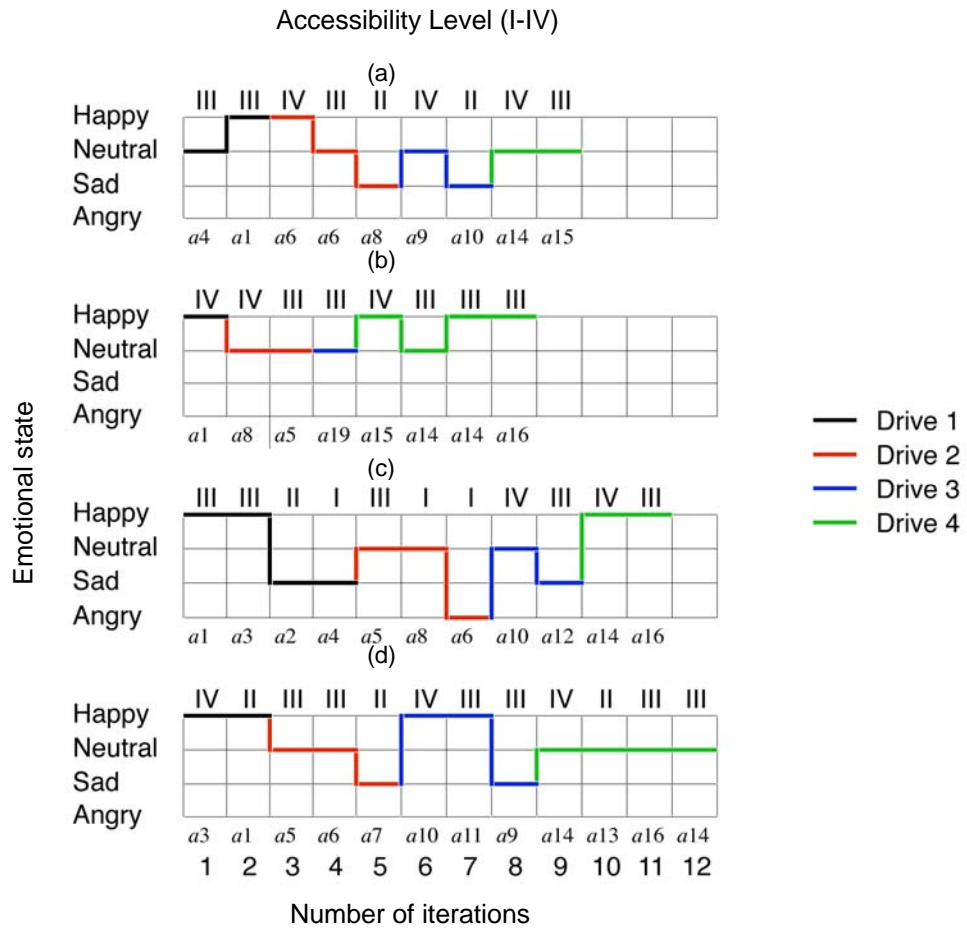


Fig 5-13: Transitions in the robot emotional state and the agent's accessibility level and the actions with online updating using the positive influence factor: (a) iterations from agent 1, (b) iterations from agent 2, (c) iterations from agent 3 and (d) iterations from agent 4.

5.2.4 Experiment #4: Implementation of Emotional State Module and Deliberative Layer with the system reliability and online updating of Markov probability matrices using the positive influence factor

Experimental Procedure

The implementation procedure of Experiment #4 is similar to Experiment #2. This experiment consists of a detailed analysis to determine the feasibility of the overall control architecture using the system reliability and a positive influence factor in the proposed assistive manner during HRI. The system reliability was used to reduce the probability of a dominant emotional state of the robot, when the robot is consecutively in this particular state and the robot is unsuccessful in satisfying its drive.

Experimental Results and Discussions

Figures 5-14- 5-17 present experimental results for experiment #4. Figure 5-14 shows the average number of iterations needed to satisfy all four drives. Figure 5-15 presents the emotional states of the robot during drive satisfaction. Figure 5-16, shows the frequency of the accessibility levels of the agent when each drive was satisfied. Figure 5-17 depicts transitions in the robot emotional state, the agent's accessibility level and the actions that were performed during drive satisfaction. Another set of experimental results is shown in Figure B-7-Figure B-9 of Appendix B.

In Figure 5-14, it took an average of 2 iterations to satisfy the robot's required drives. It can be seen in Figure 5-15 that in total, the drives were satisfied 49 percent of the time when the robot was in the emotional state of happy, 32 percent of the time when the robot was in the emotional state of neutral, 16 percent of the time when the robot was in the emotional state of sad and, 2 percent of the time when the robot was in the emotional state of angry, respectively. Figure 5-16 shows that the robot was able to satisfy the drives a significant amount of the time by utilizing all four emotional states efficiently, regardless of the accessibility level of agent. In Figure 5-17, actions *a1-a4*, *a5-a8*, *a9-a12* and *a13-a16* represent different actions for Drive 1, Drive 2, Drive 3, and Drive 4, respectively. The roman numbers (I-IV) represents accessibility levels of agent. In Figure 5-17(a) and (d), it took at most 9 iterations to satisfy all four drives. The average number of iterations that were required to satisfy all four drives was 8 iterations. The number of iterations was lower than the results in Experiment #2 and #3. In this experiment, the robot performed the appropriate actions with suitable emotional state to satisfy the drives. In Figure 5-17, the robot changed its emotional state or actions quickly to satisfy the drive when the current emotional state or action did not work. For example, in Figure 5-17(a), when drive 1 was implemented, the accessibility level transitioned from IV-I-I influencing the robot to change its emotional state from happy-sad-angry with actions *a4*, *a4*, and *a2* to satisfy the drive. This verifies that the robot has the ability of learning and adjusting its behavior based on environment changes.

Experimental results in Figure B-7 - Figure B-9 in Appendix B are similar to Figure

5-14 - Figure 5-16. It can be seen in Figure B-7, it took 2 iterations to satisfy the required drives. In Figure B-8, 55 percent, 26 percent, 18 percent and 1 percent of drives are satisfied with the emotional states of happy, neutral sad and angry, respectively. Similar to Figure 5-16, Figure B-9 shows that the robot utilized all four emotional states efficiently to satisfy the four drives, implying that the robot was able to satisfy the drives regardless the agent's accessibility level.

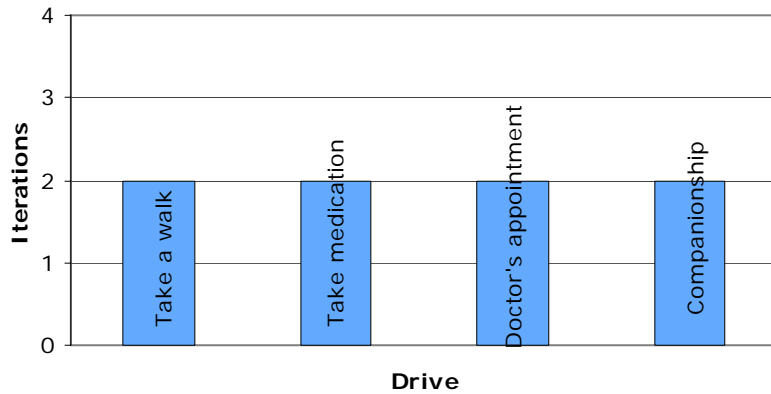


Figure 5-14: Average number of iterations needed to satisfy all four drives.

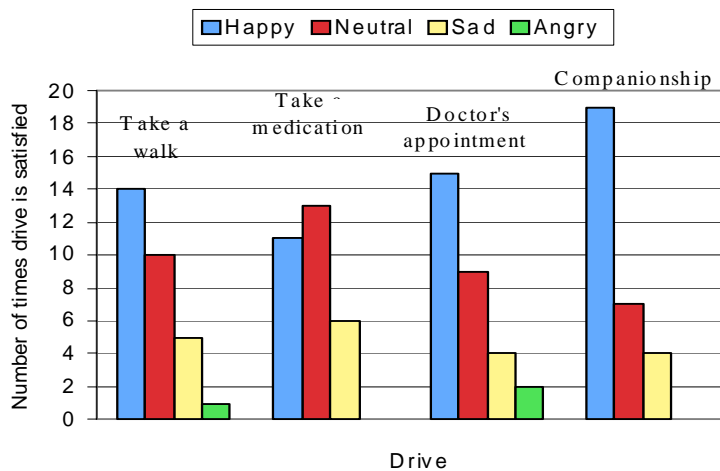
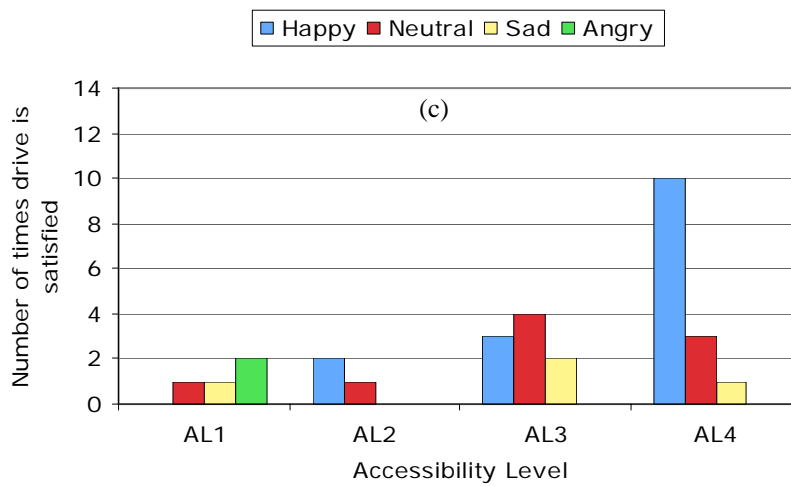
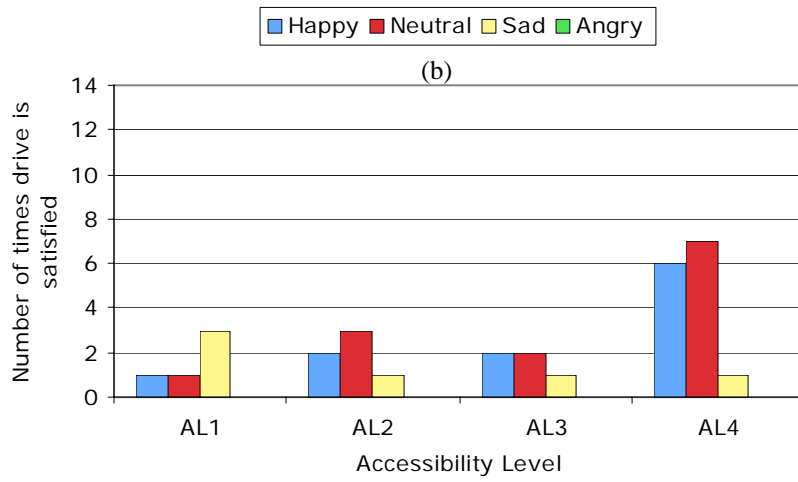
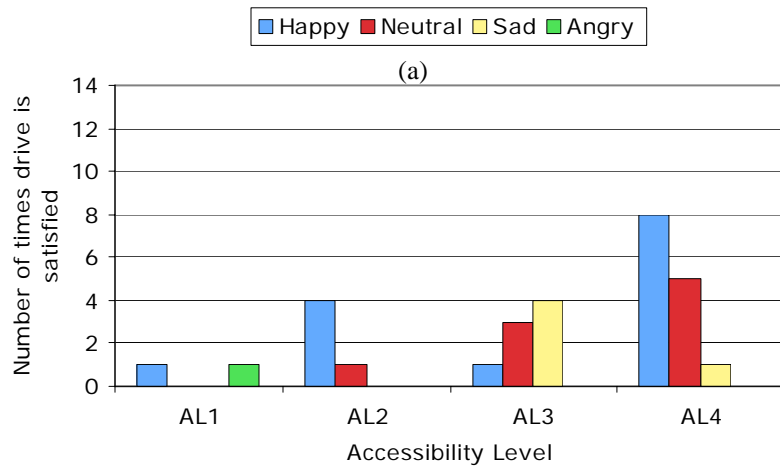


Figure 5-15: Emotional states of the robot during drive satisfaction.



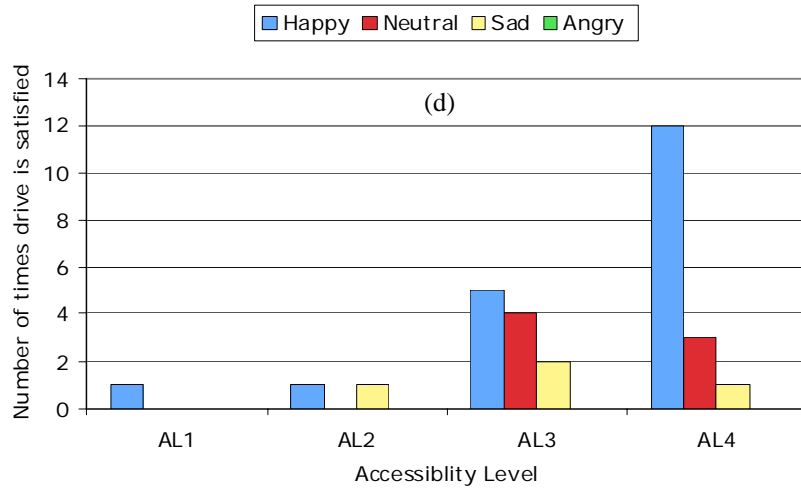


Figure 5-16: Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

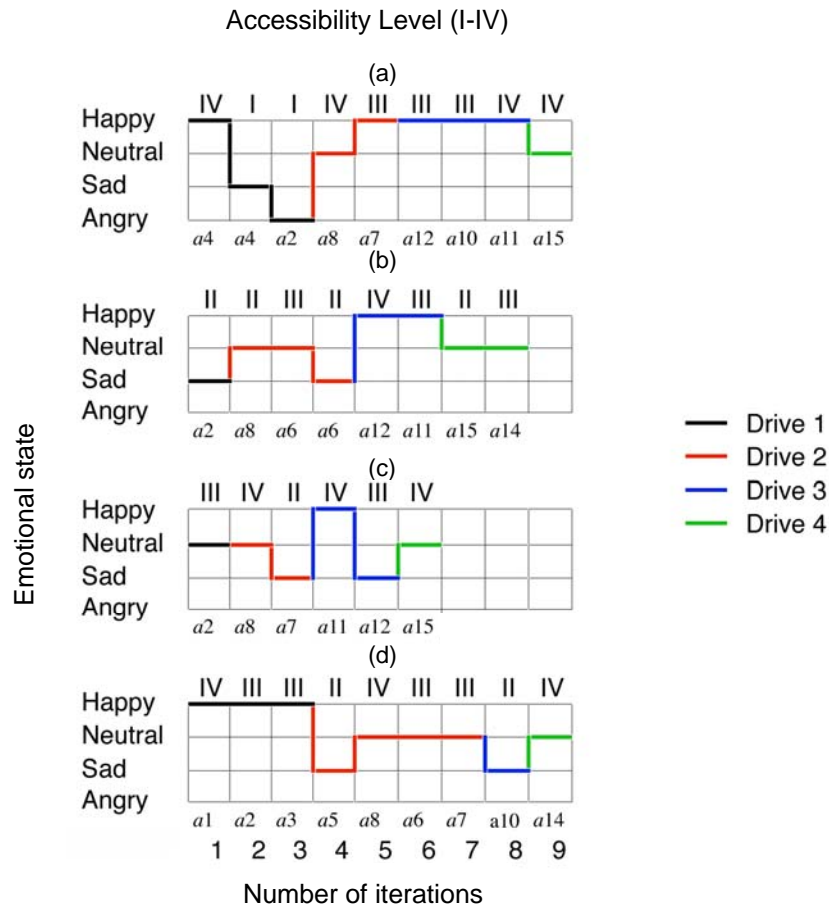


Fig 5-17: Transitions in the robot emotional state and the agent's accessibility level and the actions with the positive influence factor and the system reliability: (a) iterations from agent 1, (b) iterations from agent 2, (c) iterations from agent 3 and (d) iterations from agent 4.

Comparison and Discussion of Experiments #2, #3 and #4

Herein, three approaches were implemented to determine the robot's behavior during HRI. Of the three experiments, Experiment #4 took the least number of iterations to satisfy all four drives. Experiment #4 used 2 iterations to satisfy each drive as shown in Figure 5-14, whereas Experiment #2 satisfied all four drives with an additional 2 iterations in Figure 5-6, and Experiment #3 took an additional 1 iteration to satisfy all the drives in Figure 5-10. Although the comparison of a single set of iterations for drive completion may not seem significant, when the robot is interacting with a person, it is important to make sure that the interaction is effective. Hence, the robot should satisfy its drives with the least amount of iterations and hence minimizing the number of times the robot has to repeat itself in order to satisfy a given drive.

Overall, the approach that best supports an efficient HRI was the one implemented in Experiment #4, where both a positive influence factor and system reliability was utilized to update the emotional state module. The approach allows the robot to effectively adjust its emotional state to satisfy all four drives with a minimum number of iterations. Robot Design experiment with human subjects.

5.2.5 Experiment #5: Robot Design experiment with human subject

Experimental Procedure

Experiment #5 was carried out to test how the robot's overall control architecture was effective in real HRI scenarios with humans. In particular, two types of experiments were conducted with the robot having a real human voice and a computer simulated voice using the emotional state module and deliberative layer utilized in experiment #2. A total of 35 participants in the age group of 22-56 engaged in the interactions. Each human subject was asked questions by the robot, discussed in Section 5.1.5. The accessibility levels of the participants were determined using the DNSS. The robot is determined to satisfy a given drive when the human says "Yes" during the interaction. Otherwise, the robot continuously attempts to satisfy the drive, until the participant says "Yes". After the robot successfully satisfies all four drives for the human subject, the robot starts interacting with other human subjects.

Experimental Results and Discussions

Experimental results are shown in Figures 5-18-5-20 for the real human voice and Figures 5-21- 5-23 for the computer generated voice. Figure 5-18 and Figure 5-21 show the average number of iterations needed to satisfy all four drives. Figure 5-19 and Figure 5-21 present the emotional states of the robot during drive satisfaction. Figure 5-20 and

Figure 5-23, show the frequency of the accessibility levels of the agent when each drive was satisfied.

The results presented are similar to those presented for Experiment #2. In Figures 5-18 and 5-21, it took 2 to 3 iterations to satisfy each of the drives. In Figure 5-19, the robot was able to satisfy the four drives in the dominant emotional state of happy followed by neutral. The drives were satisfied 68 percent of the time when the robot was in the emotional state of happy, and 25 percent of the time when the robot was in the emotional state of neutral, respectively. On the other hand, in Figure 5-22, the emotional state of neutral is used to satisfy all four drives 44 percent of the time, followed by happy (39 percent). During the experiments, the participants provided feedback in which they stated they felt more comfortable talking to the robot when it speaks with a real human voice. What is interesting to note is that the robot used all four emotional states more efficiently to satisfy all the drives when the computer generated voice was utilized.

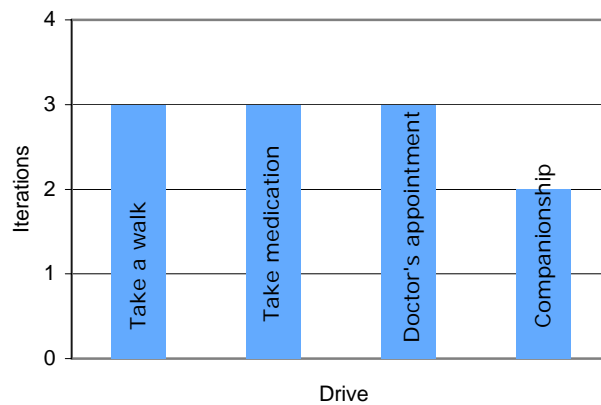


Figure 5-18: The average number of iterations needed to satisfy all four drives for the real human voice.

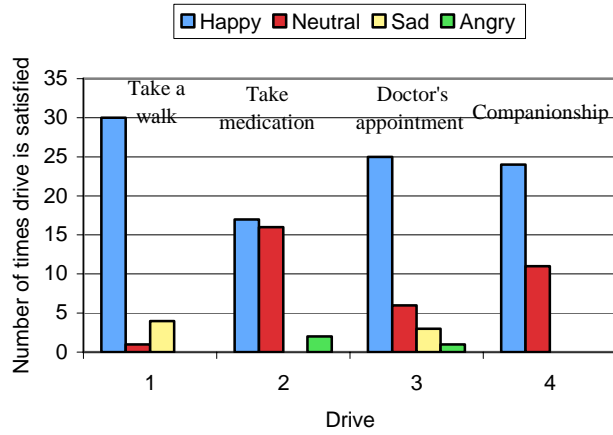
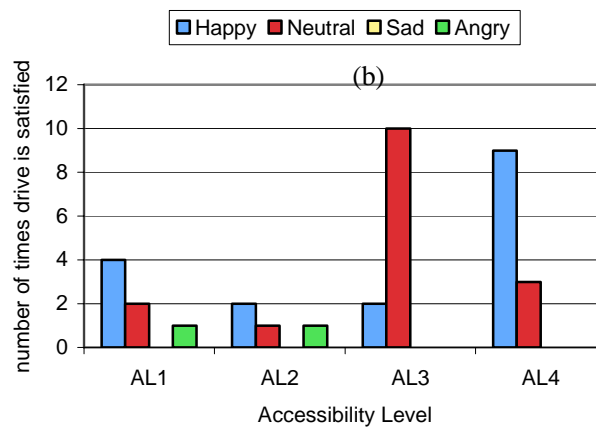
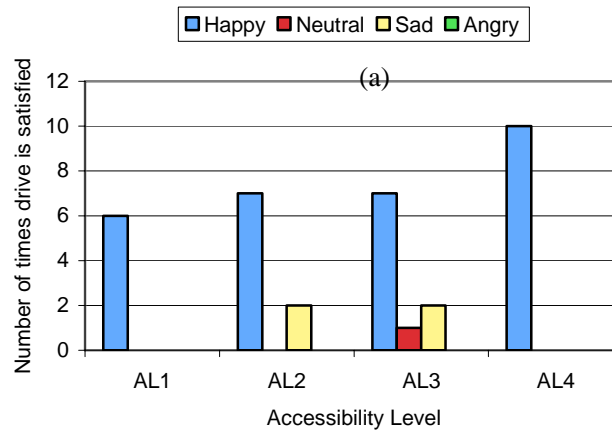


Figure 5-19: Emotional states of the robot during drive satisfaction for the real human voice.



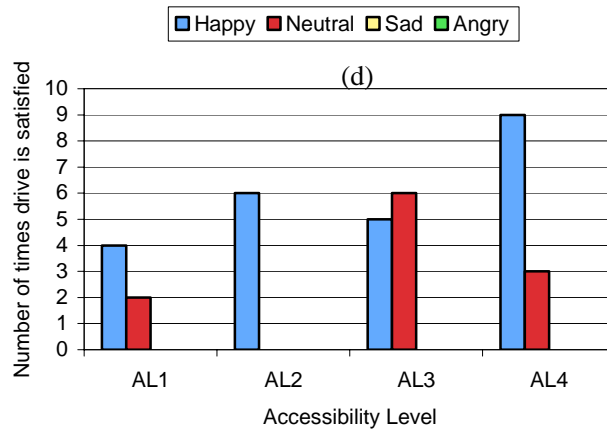
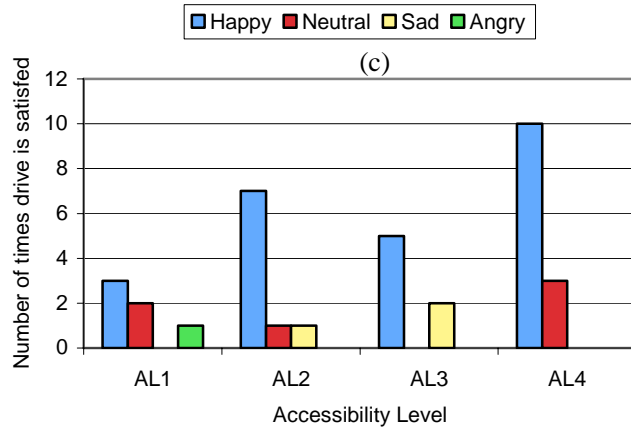


Figure 5-20: Frequency of the accessibility levels of the agent when each drive was satisfied for real human voice: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

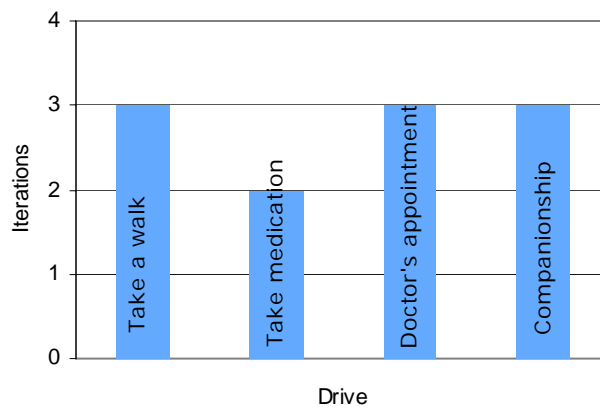


Figure 5-21: The average number of iterations needed to satisfy all four drives for the computer generated voice.

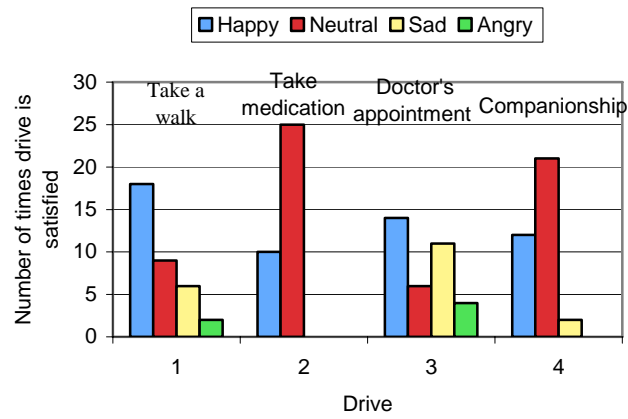
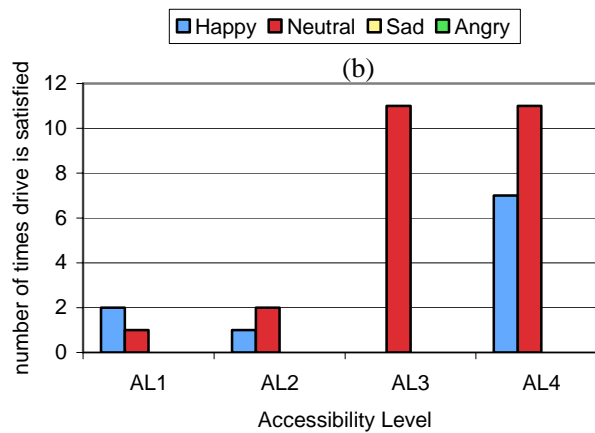
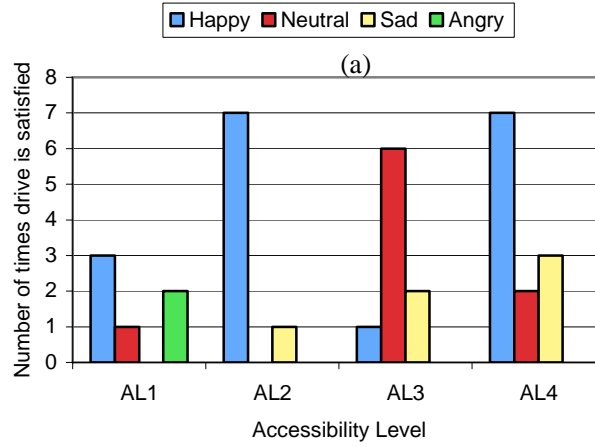


Figure 5-22: Emotional states of the robot during drive satisfaction for the generated computer voice.



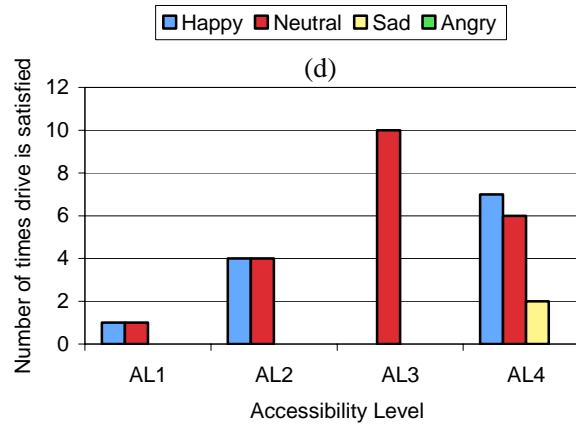
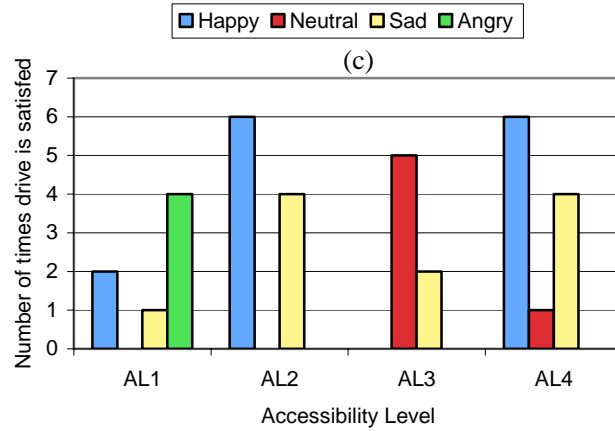


Figure 5-23: Frequency of the accessibility levels of the agent when each drive was satisfied for generated computer voice: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

5.3 Chapter Summary

This Chapter presented the results for five preliminary experiments performed to verify the effectiveness of proposed robot control modules. The proposed modules in the control architecture were integrated into a socially assistive robot. Separate experiments were performed to test the effectiveness of each control module in the HRI architecture. The experimental results verified the potential utilization of the proposed control modules in determining the appropriate robot behavior during effective HRI.

Chapter 6 Conclusions

Socially assistive robots need to have a high degree of cognitive and emotional capabilities as well as the ability to react and understand human intention during HRI. This thesis focuses on modules in a socially assistive robot that would assist in task completion. The overall objective is to design specific modules within HRI control architecture to determine the appropriate behavior of socially assistive robots during assistive interactions. In particular, two main modules utilized in determining the task-driven behavior of these types of robots are designed: (i) the emotional state module, and (ii) the deliberative layer.

6.1 Summary of Contributions

The primary contributions of this work are summarized below:

6.1.1 HRI Control Architecture

In Chapter 2, the HRI control architecture for task-driven assistive robots that was designed by our research group was presented. This thesis focuses on developing the appropriate processing mechanisms of two important modules of the HRI control architecture: (i) the robot emotional state module, and (ii) the decision making

deliberative layer. In particular, the emotional state module has been designed to be utilized in determining in real-time an assistive robot's most effective emotional state during its assistive task completion. The deliberative layer determines the explicit overall task-driven behavior of the robot, based on the robot's emotional state, the human mood state and the drives that need to be satisfied.

6.1.2 Robot Emotional State Module

In Chapter 3, the robot emotional state module within in the context of the HRI control architecture is presented. In particular, a Markov state-space representation approach is utilized to determine robot's emotional states during interaction based on the human's mood (i.e. wellbeing), the drives that need to be satisfied and the robot's current emotional state. Due to the degrading performance of a Markov chain over time, online updating of the Markov probability matrices needs to take place. In this work, a positive influence factor is introduced in order to address this problem. The positive influence factor is treated as a type of "reward" for drive satisfaction and is applied in scenarios: i) when a drive is satisfied, and ii) when a drive is not satisfied. In addition, system reliability was utilized to reduce the probability of a dominant emotional state of the robot, when the robot is consecutively in this particular state (i.e., the dominant state) and is unsuccessful in satisfying its drive. Therefore, the system reliability allows for the incorporation of the other emotional state to determine the robot's behavior.

6.1.3 Deliberative Layer

In Chapter 4, the main decision making module of the HRI control architecture is presented. In particular, a reinforcement learning method known as *Q*-learning is implemented. Since human's actions during HRI are unpredictable, a non-deterministic approach to *Q*-learning is adapted. The *Q*-learning algorithm determines the overall assistive behavior of the robot based on a set of state and action pair. The state is a function of the robot's emotional state, the human's mood state and the specific drive to satisfy and an action represents the implemented behavior of the robot.

6.1.4 Implementation

The proposed modules in the control architecture were integrated into a socially assistive robot named Brian. HRI experiments were carried out consisting of one-on-one interaction scenarios between a human and Brian in a lab setting. The robot provided reminders and companionship via four different drives in a similar manner to how it would in real-life situations. Brian is capable of being in four different emotional states: happy, neutral, sad, and angry during the interaction. The behavior of the robot is displayed by verbal and non-verbal communications. Separate experiments were performed to test the effectiveness of each control module in HRI setting. In addition, experiments were also performed to better inform the design of the robot. The experiments have shown the potential of developed control modules to provide proper interactive scenarios with appropriate emotion to accomplish the HRI.

6.2 Discussion and Future Work

The proposed learning based intelligence for a socially assistive robot developed in this thesis were first attempt at addressing two main control modules in empathizing on human's wellbeing. Hence, any future work in this research area should improve algorithm of processing modules that have been developed. There are some specific issues that may strengthen the results of this work.

The socially assistive robot is able to learn the environment and implement the appropriate actions through exploration and exploitation of the environment. However, in some situations, the robot tends to implement the same actions in consecutive times. In this thesis, the exploration of the environment for the robot was implemented in off-line process and was not performed in on-line process. As a result, a Q -value of specific action becomes significantly higher than other actions and the action is chosen consecutively. Hence, the integrations of an algorithm that can perform the on-line exploitation during the exploration process will need to be considered.

The proposed control modules are limited to the number of drives and actions that can be implemented, as well as the robot's emotional state. For future work, a methodology that can efficiently process large number of drives, actions and the robot emotional state need to be considered. Moreover, processing for the remaining modules in the overall HRI control architecture will need to be developed. In addition, the controller will be expanded to incorporate multi-modal inputs from the human, i.e.,

speech and facial expressions in addition to gestures. Once all the designs are optimized, the socially assistive robot can be tested in actual health care environment for its effectiveness.

6.3 Final Concluding Statement

In overall conclusion, the recent research efforts on the development of control module for robotic task-driven behavior provides the research community with insight into how to design learning based intelligence for robots involved in real-time human-robot interaction. Future work in this research area will help in the implementation to health care applications.

References

- [1] K. Kinsella, V.A. Velkolt, "An aging world-2001", U.S Department of Health and Human Services and Human Services and U.S Census Bureau Report.
- [2] R. Maria Li, A.C. Ladarola, and C. Maisano, "Why population aging matters- A global perspective", National Institute on Aging and U.S. Department of Health and Human Services Report, Available HTTP: <http://www.state.gov/g/oes/rls/or/81537.htm#over>
- [3] United States Accountability Offices Report, Washington D.C, 4-7, <http://www.gao.gov/new.items/d07492r.pdf>, 2007.
- [4] P. Zurn, M. Poz, B. Stilwell and O. Adams, "World health organization report", Imbalances in the Health Workforce, 2002.
- [5] M. Nies, G. Nejat, Y. Sun, and T. Sexton, "A multidisciplinary team program of research with interactive nurse robots to enhance elderly health", Sigma Theta Tau International Nursing Research Congress Focusing on Evidence-based Practice, Vienna, 2007.
- [6] The Center for Health Program Development and Management, "Working paper: Nurses' Workplace Issues", Patient Safety, and the Quality of Patient Care, 2001.

- [7] K. Freedman, S. Mataric, M.J. Cunningham, and M.J. Lopez, "A hands-off physical therapy assistance robot for cardiac patients", IEEE International Conference on Rehabilitation Robotics, Chicago, 2005, pp.337-340.
- [8] R.M. Mahoney, H.F. Machiel Van der Loos, P.S. Lum, and C. Burgar, "Robotic stroke therapy assistant", *Robotica* 21 (2003) 33-44.
- [9] T.Nef, R. Riener, "ARMin – design of a novel arm rehabilitation robot", IEEE International Conference on Rehabilitation Robotics, Chicago, 2005, pp.57-60.
- [10] H. I. Krebs, N. Hogan, M. L. Aisen, and B. T. Volpe, "Robot-aided neurorehabilitation", *IEEE Transactions on Rehabilitation Engineering* 6 (1) (1998)75-87.
- [11] P. S. Lum, C. G. Burgar, and P. C. Shor, "Evidence for improved muscle activation patterns after retraining of reaching movements with the MIME robotic system in subjects with post-stroke hemiparesis", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*(2)(2004)186-194.
- [12] S. Lum, C. G. Burgar, and P. C. Shor, "Use of the MIME robotic system to retrain multijoint reaching in post-stroke hemiparesis: why some movement patterns work better than others", IEEE International Conference of the Engineering in Medicine and Biology Society, Richmond, 2003, pp. 1475- 1478.
- [13] R. Riener, T. Nef, and G. Colombo, "Robot-aided neurorehabilitation of the upper extremities", *Medical and Biological Engineering and Computing* 43(2005)2-10.
- [14] S. Dubowsky, "PAMM - a robotic aid to the elderly for mobility assistance and monitoring: A helping-hand for the elderly", IEEE International Conference Robotics and Automation, San Francisco, 2000, pp. 570-576.

- [15] Peter Aigner, Brennan J. McCarragher, “Shared control framework applied to a robotic aid for the blind”, IEEE International Conference on Robotics and Automation, Leuven, 1998, pp.717-722.
- [16] J. Glover, D. Holstius, M. Manojlovich, K. Montgomery, A. Powers, J. Wu, S. Kiesler, J. Matthews, and S. Thrun, “A robotically-augmented walker for older adults, Technical report”, Carnegie Mellon University, School of Computer Science, 2003.
- [17] S.P. Levine, D.A. Bell, L.A. Jaros, R.C. Simpson, Y. Koren, J. Borenstein, “The NavChair assistive wheelchair navigation system”, IEEE Transition on Rehabilitation Engineering 7 (4) (1997) 443-451.
- [18] C. Galindo, J. González, and J.-A. Fernandez-Madrigal, “A control architecture for human-robot integration application to a robotic wheelchair”, IEEE Transactions on Systems, Man, and Cybernetics 36(6)(2006)1053-1068.
- [19] B. J. Driessen, H. G. Evers, J. A. van Woerden, “MANUS--a wheelchair-mounted rehabilitation robot”, Institute of Mechanical Engineers 215(3) (2001)285-290.
- [20] P. Aigner, B. McCarragher, “Shared control framework applied to a robotic aid for the blind”, IEEE International Conference on Robotics and Automation, Canberra, 1998, pp. 717-722. 1998.
- [21] S. Butner, M. Ghodoussi, “Transforming a surgical robot for human telesurgery”, IEEE Transition on Robotics and Automation 19(5) (2003)818–824.
- [22] Intuitive Surgical, Available HTTP: <http://www.intuitivesurgical.com>
- [23] U. Hagn, M. Nickl, S. Jorg, G. Passig, T. Bahls, A. Nothhelfer, F. Hacker, L. Le-Tien, A. Albu-Schäffer, R. Konietschke, M. Grebenstein, R. Warpup, R. Haslinger, M.

- Frommberger, G. Hirzinger, “The DLR MIRO: a versatile lightweight robot for surgical applications”, *Industrial Robot: An International Journal* 35(2008)324-336.
- [24] S. Peddatham, W. Peine, H. Z. Tan, “Assessment of vibrotactile feedback in a needle-insertion task using a surgical robot”, *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Reno, 2008*, pp. 93-99.
- [25] ScriptPro, Available HTTP: <http://www.scriptpro.com/products/sp-200/main.htm>
- [26] Mckesson, Available HTTP: http://www.mckesson.com/en_us/McKesson.com/
- [27] Parata, Available HTTP: <http://www.parata.com/fill/rds.php>
- [28] AmerisourceBergen, Available HTTP: <http://www.amerisourcebergen.com/cp/1/>
- [29] W. Bantas, M. Wolfram, R Bräutigam, M. Probst, W. Beecken, D. Jonas, J. Binder, “Da Vinci robot assisted Anderson-Hynes dismembered pyeloplasty: technique and 1 year follow-up”, *World Journal of Urology* 21(3) (2003) 133-138
- [30] D, Feil-Seifer, M. Mataric, “Defining socially assistive robotics”, In *Proceedings of the International Conference on Rehabilitation Robotics. Chicago, IL; 2005:465-468*
- [31] T. Shibata, “An overview of human interactive robots for psychological enrichment”, *Proceedings of the IEEE* 92 (11) (2004)1749-1757.
- [32] M. Kanoh, S. Kato, and H. Itoh, “Emotive facial expressions using emotional space in sensitivity communication robot “ifbot””, *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004*, pp. 1586-1591.
- [33] University of Hertfordshire in UK, Adaptive systems research group and robotcup project in Europe, KASPAR: Kinesics and Synchronization in Personal Assistant Robotics <http://kaspar.feis.herts.ac.uk/>.

- [34] M. Montemerlo, J. Pineau, N. Roy, S. Thrun and V. Verma, “Experiences with a mobile robotic guide for the elderly”, National Conference on Artificial Intelligence, Edmonton, 2002, pp. 587-592.
- [35] S. Freedman, M.J. Mataric, M.J. Cunningham and B. Lopez, “A hands-off physical therapy assistance robot for cardiac patients”, IEEE International Conference on Rehabilitation Robotics, Chicago, 2005, pp. 337-340.
- [36] K. Sato, M. Ishii, and H. Madokoro, “Testing and evaluation of a patrol robot system for hospitals”, Electronics and Communication in Japan 86 (12) (2003)14 – 26.
- [37] M. Lopez, R. Barea, L. Bergasa and M. Escudero, “A human–robot cooperative learning system for easy installation of assistant robots in new working Environments”, Journal of Intelligent and Robotic System 40(3)(2004)233-265.
- [38] A. Libin and E. Libin, “Person–robot interaction from the robopsychologists point of view: The Robotic Psychology and Robotherapy Approach”, Proceedings of the IEEE 92 (11) (2004) 1789–1803.
- [39] T. Hashimoto, M. Senda, and H. Kobayashi, “Realization of realistic and rich facial expressions by face robot”, IEEE Technical Exhibition Based Conference on Robotics and Automation, Tokyo, 2004, pp. 37-38.
- [40] K. Hayashi, Y. Onishi, K. Itoh, H. Miwa, and A. Takanishi, “Development and evaluation of face robot to express various face shape”, IEEE International Conference on Robotics and Automation, Orlando, 2006, pp. 481-486.
- [41] W. Weiguo, and M. Qingmei, “Development of the humanoid head portrait robot system with flexible face and expression”, IEEE International Conference on Robotics and Biomimetics, Shenyang, 2004, pp. 718-723.

- [42] A. Sloman, "Varieties of affect and the CogAff architecture schema", AISB'01 Symposium on Emotion, Cognition and Affective Computing, York, 2001, pp. 39-48.
- [43] H.R. Kim, K.W. Lee, and D.S. Kwon, "Emotional interaction model for a service robot", IEEE International Workshop on Robots and Human Interactive Communication, Nashville, 2005, pp. 672-678.
- [44] R. Barber and M.A. Salichs, "A new human based architecture for intelligent autonomous robots", The Fourth IFAC Symposium on Intelligent Autonomous Vehicles, Sapporo, 2001, pp. 85-90.
- [45] M. Malfaz and M.A. Salichs, "A new architecture for autonomous robots based on emotions", IFAC Symposium on Intelligent Autonomous Vehicles, Lisbon, 2004.
- [46] J.D. Velasquez, "Modeling emotions and other motivations in synthetic agents", National Conference on Artificial Intelligence, Providence, 1997, pp. 10-15.
- [47] J.D. Velasquez, "From affect programs to higher cognitive emotion: an emotion-based control approach", Workshop on emotion-based agent, 1999, pp114-120.
- [48] C. Breazeal, B. Scassellati, "Infant-like social interactions between a robot and a human caretaker", *Adaptive Behavior* 8(1)(2000) 49-74.
- [49] C. Breazeal and B. Scassellati, "A context-dependent attention system for a social robot", International Joint Conference on Artificial Intelligence, Stockholm, 1999, pp. 1146-1153.
- [50] J. Hirth, T. Bram, K. Berns, "Emotion based control architecture for robotics applications", *Advances in Artificial Intelligence* 4667(2007) 464-467.

- [51] K. Wada, T. Shibata, T. Saito, K. Tanie, "Robot assisted activity for elderly people and nurses at a day service center", IEEE International Conference on Robotics and Automation, Washington, 2002, pp. 1416-1421.
- [52] C. Park, J. W. Ryu, S. Kang, J. Kim, J. Sohn, H. Cho, "The emotion expression robot through the affective interaction: KOBIE", Proceedings of the 1st international conference on Robot communication and coordination 318(2007) 53.
- [53] J. Terao, M. Ficocelli, G. Nejat, "Promoting natural interactions between human and robots- the design of an intelligent socially assistive robot", Robotics and Autonomous system, under review.
- [54] G. Nejat and M. Ficocelli, "Can I be of assistance? The intelligence behind an assistive robot", IEEE International Conference on Robotics and Automation, Pasadena, 2008, pp. 3564-3569.
- [55] G. Nejat, J. Terao, L. Trejos, Z. Zhang, "The design of an intelligent socially assistive robot for elderly care", ASME International Mechanical Engineering Congress and Exposition, Boston, 2008, IMECE2008-67678.
- [56] M. Davis, and D. Hadiks, "Nonverbal behavior and client state changes during psychotherapy", Journal of Clinical Psychology 46 (3) (1990) 340-35.
- [57] M. A. Arbib and J. M. Fellous, "Emotions: From brain to robot", Trends in Cognitive Science. Elsevier Science, 2004.
- [58] P. Ekman, "An argument for basic emotions", Cognition and Emotion 6 (3) (1992) 169-200.
- [59] D. Evans, "Can robots have emotions?", Psychology Review 11 (1)(2003) 2-5.

- [60] A. Sloman, M. Croucher, "Why robots will have emotions", In Proceedings 7th International Joint Conference on AI, Vancouver, 1981, pp. 197-202.
- [61] R. W. Picard, "What does it mean for a computer to "have" emotions?", MIT Medialab Report Technical Report 534, 2001.
- [62] H. Miwa, T. Okuchi, K. Itoh, H. Takanobu, A. Takanishi, "A new mental model for humanoid robots for human friendly communication - introduction of learning system", mood vector and second order equations of emotion, IEEE International Conference on Robotics & Automation, Taipei, 2003, pp. 3588-3593.
- [63] H. Miwa, K. Itoh, M. Matsumoto, M. Zecca, H. Takanobu, S. Roccella, M.C. Carrozza, P. Dario, A. Takanishi, "Effective emotional expressions with emotion expression humanoid robot WE-4RII", IEEE/RSJ International Conference on Intelligent Robots and Systems, Submitted, 2004, pp. 2203-2208.
- [64] G. A. Hollinger, Y. Georgiev, A. Manfredi, Z. A. Pezzementi, B. Mitchell, and B. A. Maxwell, "Design of a social mobile robot using emotion-based decision mechanisms", IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, 2006.
- [65] T. Fukuda, "An intelligent robotic system based on a fuzzy approach", Proceedings of the IEEE 87(9) (1999) 1448-1470.
- [66] N. Kubota, "Evolving pet robot with emotional model", IEEE Congress on Evolutionary Computation, La Jolla, 2000, pp. 1231-1237.
- [67] T. Y. Choi, J. Y. Lee, J. H. Shin, J. J. Lee, "Modeling of the emotional model with friendship for familiarity of robot", IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005, pp. 1235- 1240.

- [68] A. Chandra, "A computational architecture to model human emotions", IEEE Proceedings of the Intelligent Information Systems, Grand Bahama Island, 2007, pp. 86-89.
- [69] K. Kuhnlenz and M. Buss, "Towards an emotion core based on a hidden markov model", IEEE International Workshop on Robot and Human Interactive Communication, Munchen, 2004, pp. 119-124.
- [70] D. Yu, Y. Tang, J. Fang, Y. Zhou and M. Sun, "Homogeneous markov chain for modeling emotional interactions", IEEE International Conference on Advanced Communication Technology, Vol 1, Gangwon-Do, 2008, pp. 265-269.
- [71] D.A. Dawson, "Introduction to Markov Chains", Canadian Mathematical Congress, Canada, 1970.
- [72] D. Skulj, "Finite discrete time markov chains with interval probabilities", Advances in Soft Computing 37 (2006) 299-306.
- [73] R.C. Cheung, "A user-oriented software reliability model", IEEE Transition Software Engineering 6 (2)(1980) 118-125.
- [74] K. Seigrist, "Reliability of systems with markov transfer of control", IEEE Transition Software Engineering 14 (7)(1988)1049-1053.
- [75] K. Seigrist, "Reliability of systems with markov transfer of control II", IEEE Transition Software Engineering 14 (10) (1988) 1478-1480.
- [76] S. S. Gokhale and K. S. Trivedi, "Reliability prediction and sensitivity analysis based on software architecture", IEEE International Symposium on Software reliability Engineering, Connecticut, 2002, pp. 64-75.

- [77] L. Zu, P. Yang, L. Chen, X. Zhang, Y. Tian, "Obstacle avoidance of multi mobile robots based on behavior decomposition reinforcement learning", IEEE International Conference on Robotics and Biomimetics, Sanya, 2007, pp. 1018-1023.
- [78] J. Baxter, A. Tridgell, L. Weaver, "Learning to play chess using temporal differences", Machine Learning 40 (3)(2000)243-263.
- [79] T. M. Mitechell, "Machine Learning", McGraw Hill, 1997.
- [80] A. Geramifard, M.H. Bowling, R. S. Sutton, "Incremental Least-Squares, Temporal difference learning", Association for the Advancement of Artificial Intelligence 2006 pages 356-361.
- [81] K. Ishikawa, A. Sakurai, T. Fujinami, S. Kunifuji, "R-learning with multiple state-action value tables", Electrical Engineering in Japan 159(3)(2007)34-47.
- [82] D.G. Shapiro, P. Langley, R. D. Shachter, "Using background knowledge to speed reinforcement learning in physical agents", Proceedings of the 5th International Conference on Autonomous Agents, Montreal, 2001, pp. 254-261.
- [83] K. M. Gupta, "Performance comparison of sarsa(λ) and Watkin's $Q(\lambda)$ algorithms", Department of Computer Science in Texas Tech University.
- [84] C. Henrique, C. Ribeiro, "A tutorial on reinforcement learning techniques", In Proceedings of International Conference on Neural Networks, Washington, 1999.
- [85] Y. Li, C. Li, Z. Zhang, " Q -learning based method of adaptive path planning for mobile robot", IEEE International Conference on Information Acquisition, Shandong, 2006, pp. 983-987.

- [86] W.D. Smart, and L.P. Kaelbling, “Effective reinforcement learning for mobile robots”, IEEE International Conference on Robotics and Automation, Washington, 2002, pp. 3404-3420.
- [87] M. S. Park, “Intelligent transportation system using q-Learning”, IEEE International Conference on Systems, Man and Cybernetics, Washington, 2003, pp. 4684- 4687
- C.Watkins and P.Dayan, Machine learning: Q-learning, Machine Learning 8 (1992) 279-292.
- [88] B. Allison and G. Nejat, An expressive socially assistive robot for health care applications, ASME Design Engineering Technical Conferences & Computers and Information in Engineering Conference, New York City, 2008, DECT2008-50082.
- [89] B. Allison, G. Nejat, and E. Kao, The design of an expressive human-like socially assistive robot, ASME Journal on Mechanisms and Robotics, In print.
- [90] G. Nejat, B. Allison, N. Gomez, and A. Rosenfeld, The design of an interactive socially assistive robot for patient care, ASME International Mechanical Engineering Congress and Exposition, Seattle, 2007, IMECE2007-41811.

Appendix A

The following example show how the state reliability function work.

$$\hat{A} = \begin{matrix} & e_H & e_N & e_S & e_A \\ e_H & 17 & 7 & 6 & 5 \\ e_N & 11 & 10 & 3 & 6 \\ e_S & 11 & 4 & 7 & 1 \\ e_A & 7 & 5 & 4 & 9 \end{matrix}, \hat{B}_1 = \begin{matrix} & x_{H1} & x_{H2} & x_{H3} & x_{H4} \\ e_H & 13 & 7 & 6 & 16 \\ e_N & 3 & 7 & 8 & 9 \\ e_S & 5 & 9 & 1 & 6 \\ e_A & 11 & 6 & 1 & 5 \end{matrix}, \hat{B}_2 = \begin{matrix} & d_1 & d_2 & d_3 & d_4 \\ e_H & 8 & 10 & 11 & 11 \\ e_N & 8 & 7 & 7 & 11 \\ e_S & 8 & 4 & 4 & 6 \\ e_A & 5 & 7 & 6 & 0 \end{matrix}$$

$$A = \begin{bmatrix} 0.369 & 0.269 & 0.3 & 0.238 \\ 0.239 & 0.384 & 0.15 & 0.285 \\ 0.239 & 0.153 & 0.35 & 0.047 \\ 0.152 & 0.192 & 0.2 & 0.428 \end{bmatrix}, B_1 = \begin{bmatrix} 0.40625 & 0.241 & 0.37 & 0.444 \\ 0.09375 & 0.241 & 0.5 & 0.25 \\ 0.15625 & 0.310 & 0.06 & 0.168 \\ 0.34375 & 0.206 & 0.06 & 0.1389 \end{bmatrix}, B_2 = \begin{bmatrix} 0.275 & 0.357 & 0.392 & 0.392 \\ 0.275 & 0.25 & 0.25 & 0.392 \\ 0.275 & 0.142 & 0.142 & 0.214 \\ 0.175 & 0.25 & 0.214 & 0.000 \end{bmatrix}$$

Assume previous robot emotional state, $x_R(k-1)=e_H$ =happy, current drive, $d(k)$ =drive 1 and current human's accessibility level, $x_H(k)$ = level 3.

By using Equation (3-1) and (3-2),

$$x_R(k) = \begin{bmatrix} 0.369 \\ 0.239 \\ 0.239 \\ 0.152 \end{bmatrix} * \begin{bmatrix} 0.37 \\ 0.5 \\ 0.06 \\ 0.06 \end{bmatrix} * \begin{bmatrix} 0.275 \\ 0.275 \\ 0.275 \\ 0.175 \end{bmatrix} = \begin{matrix} e_H & 0.0375 \\ e_N & 0.0328 \\ e_S & 0.0039 \\ e_A & 0.0016 \end{matrix}$$

Since 0.0375 is the dominant number, the current robot emotional state $x_R(k)$ is determined as happy.

The emotional state of happy occurred 2 times in a row.

Therefore, the system reliability, Equation (3-15), is calculated as $r_i(k) = \frac{17-1}{17}$ and applied

to element of happy of probability matrix A.

$$x_R(k) = \begin{bmatrix} 0.369 * \frac{17}{17+1} \\ 0.239 \\ 0.239 \\ 0.152 \end{bmatrix} * \begin{bmatrix} 0.37 \\ 0.5 \\ 0.06 \\ 0.06 \end{bmatrix} * \begin{bmatrix} 0.275 \\ 0.275 \\ 0.275 \\ 0.175 \end{bmatrix} = \begin{bmatrix} 0.0354 \\ 0.0328 \\ 0.0039 \\ 0.0016 \end{bmatrix}$$

After the system reliability is applied, 0.0354 - the happy emotional state, will then be displayed as the dominant robot emotional state at time step k .

If this iteration had not satisfied drive 1, the system reliability, Equation (3-14),

$\varepsilon_i(k+1) = \frac{1}{(m-1)} = \frac{1}{4-1}$ is applied to the remaining elements of \hat{A} , \hat{B}_1 and \hat{B}_2 .

These are the new matrixes.

$$\hat{A} = \begin{bmatrix} 17 & 7 & 6 & 5 \\ 11.333 & 10 & 3 & 6 \\ 11.333 & 4 & 7 & 1 \\ 7.333 & 5 & 4 & 9 \end{bmatrix}, \hat{B}_1 = \begin{bmatrix} 13 & 7 & 6 & 16 \\ 3 & 7 & 8.333 & 9 \\ 5 & 9 & 1.333 & 6 \\ 11 & 6 & 1.333 & 5 \end{bmatrix}, \hat{B}_2 = \begin{bmatrix} 8 & 10 & 11 & 11 \\ 8.333 & 7 & 7 & 11 \\ 8.333 & 4 & 4 & 6 \\ 5.333 & 7 & 6 & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 0.361 & 0.269 & 0.3 & 0.238 \\ 0.241 & 0.384 & 0.15 & 0.285 \\ 0.241 & 0.153 & 0.35 & 0.047 \\ 0.156 & 0.192 & 0.2 & 0.428 \end{bmatrix}, B_1 = \begin{bmatrix} 0.40625 & 0.241 & 0.35 & 0.444 \\ 0.09375 & 0.241 & 0.49 & 0.25 \\ 0.15625 & 0.310 & 0.078 & 0.168 \\ 0.34375 & 0.206 & 0.078 & 0.1389 \end{bmatrix}, B_2 = \begin{bmatrix} 0.267 & 0.357 & 0.392 & 0.392 \\ 0.278 & 0.25 & 0.25 & 0.392 \\ 0.278 & 0.142 & 0.142 & 0.214 \\ 0.178 & 0.25 & 0.214 & 0.000 \end{bmatrix}$$

The robot emotional state, $x_R(k) = e_H = \text{happy}$ and the current drive, $d(k+1) = \text{drive 1}$ and the current human accessibility level $x_H(k+1) = \text{level 3}$.

By using Equation (3-1) and (3-2),

$$x_R(k+1) = \begin{bmatrix} 0.361 \\ 0.241 \\ 0.241 \\ 0.152 \end{bmatrix} * \begin{bmatrix} 0.35 \\ 0.49 \\ 0.0078 \\ 0.0078 \end{bmatrix} * \begin{bmatrix} 0.267 \\ 0.278 \\ 0.278 \\ 0.178 \end{bmatrix} = \begin{bmatrix} 0.0337 \\ 0.0329 \\ 0.0005 \\ 0.0020 \end{bmatrix},$$

happy is the dominant emotion again.

This is the second time the emotional state of happy failed to satisfy the drive.

Therefore, the system reliability, is calculated as $r_i(k+1) = \frac{17-2}{17}$ and applied to element of

happy of probability matrix A.

$$x_R(k+1) = \begin{bmatrix} 0.361 * \frac{17}{17+2} \\ 0.241 \\ 0.241 \\ 0.152 \end{bmatrix} * \begin{bmatrix} 0.35 \\ 0.49 \\ 0.0078 \\ 0.0078 \end{bmatrix} * \begin{bmatrix} 0.267 \\ 0.278 \\ 0.278 \\ 0.178 \end{bmatrix} = \begin{bmatrix} 0.0301 \\ 0.0329 \\ 0.0005 \\ 0.0020 \end{bmatrix}$$

After the system reliability is applied, the neutral emotional state, 0.03529, is then displayed as the dominant robot emotional state at time step $k+1$.

Appendix B

Figure B-1 – Figure B-3 show experimental results for Experiment #3 with no updating method for the robot emotion model. 30 participants were utilized in the experiment.

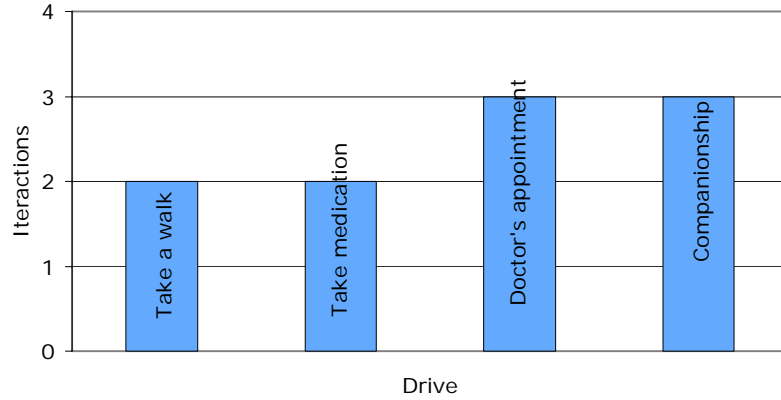


Figure B-1: Average number of iterations needed to satisfy all four drives.

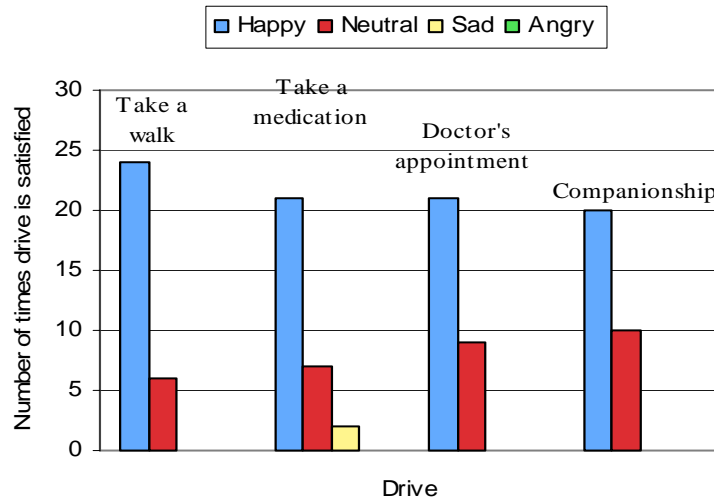
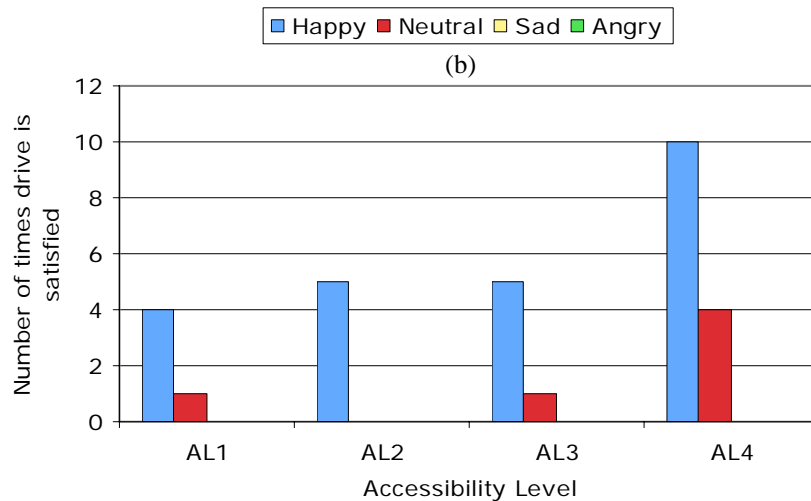


Figure B-2: Emotional states of the robot during drive satisfaction.



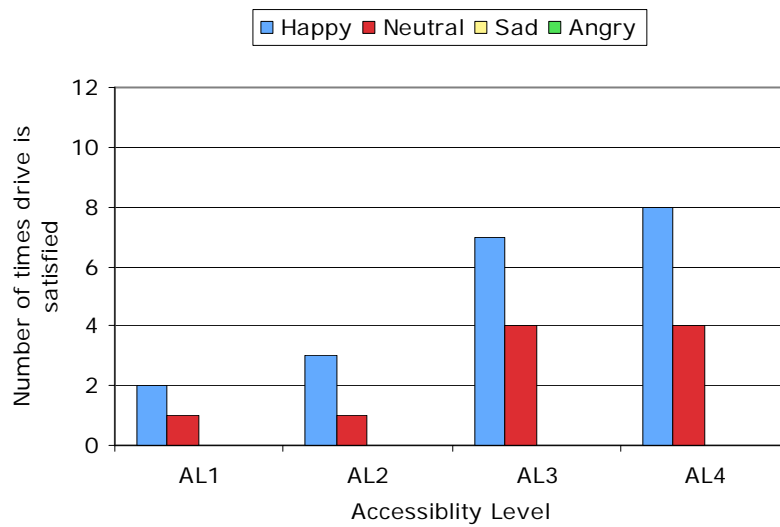
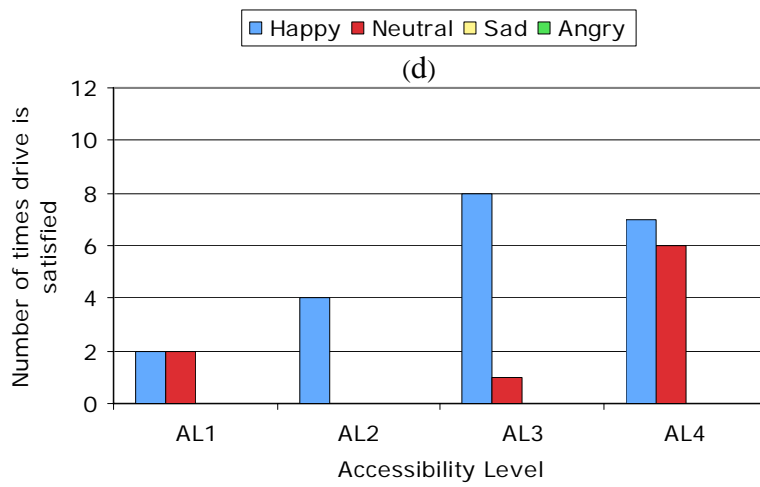
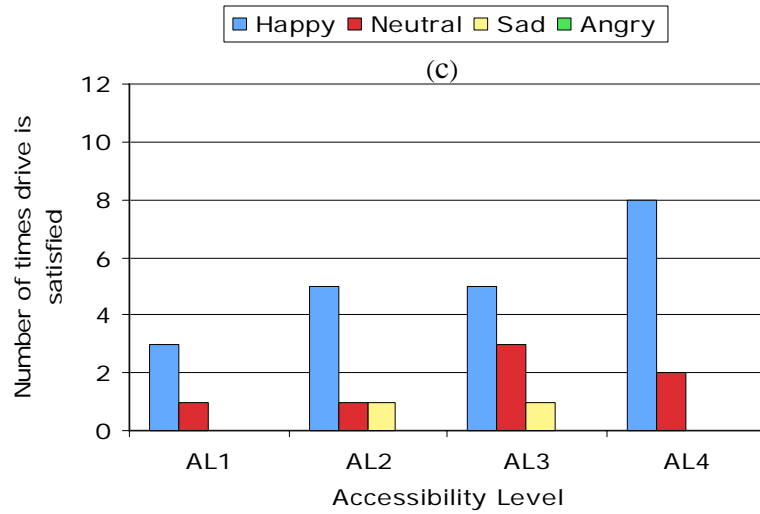


Figure B-3: Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

Figure B-4 – Figure B-6 show experimental results for Experiment #4 with the positive influence factor for the robot emotion model. 30 participants were utilized in the experiment.

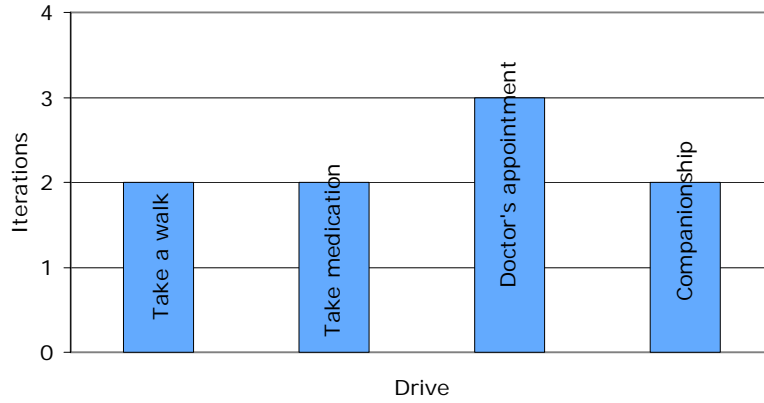


Figure B-4: Average number of iterations needed to satisfy all four drives.

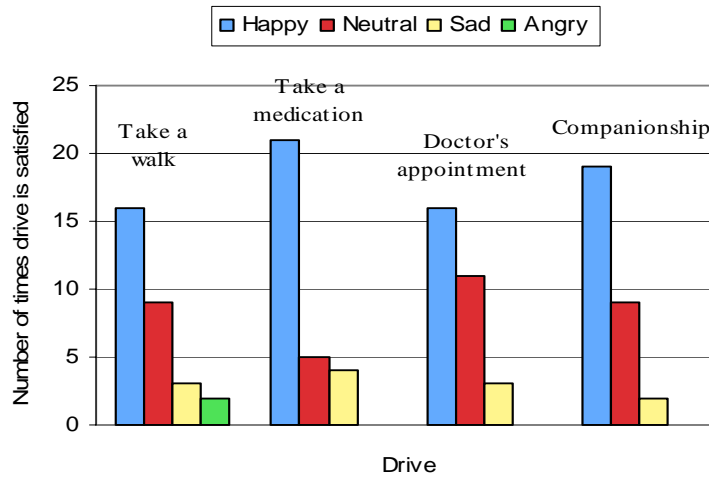
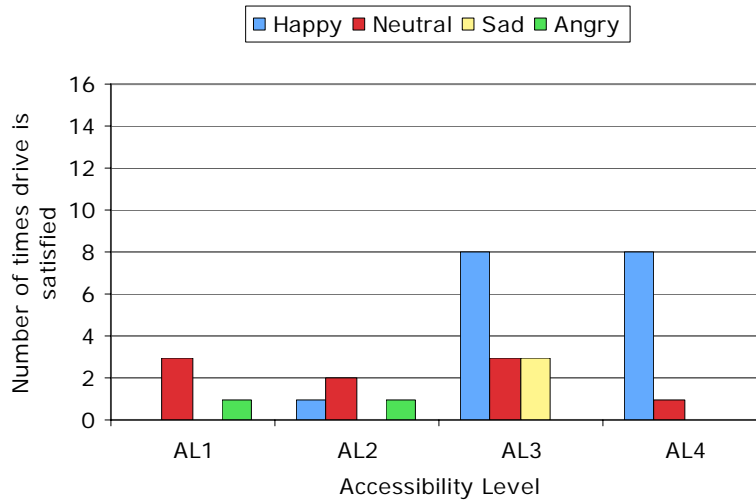


Figure B-5: Emotional states of the robot during drive satisfaction.



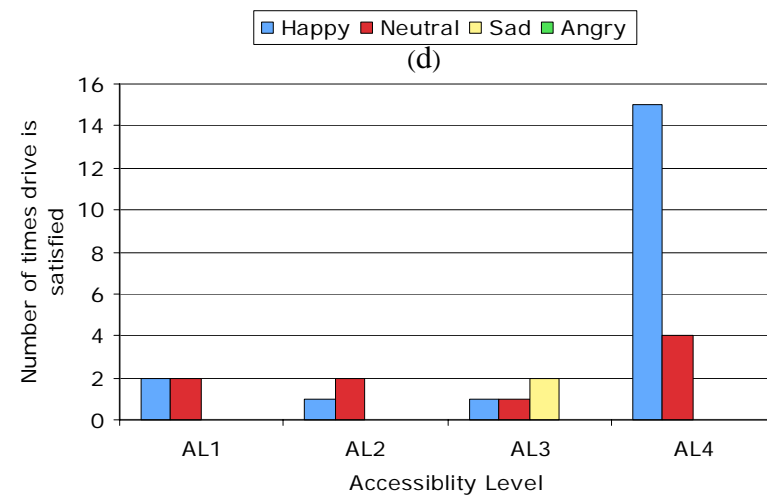
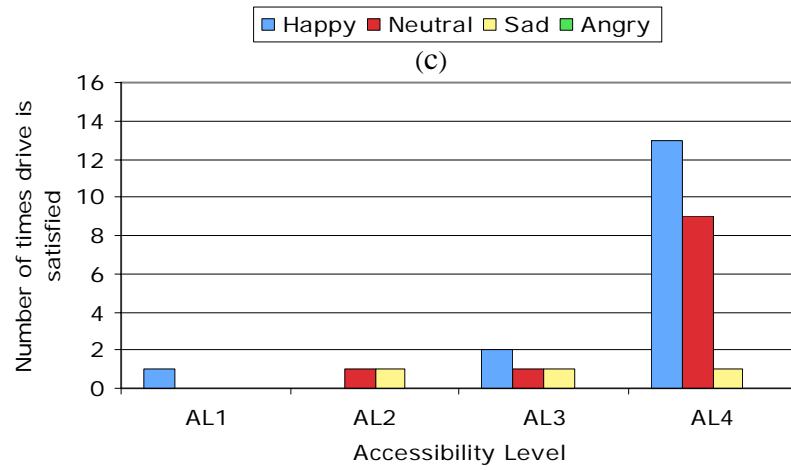
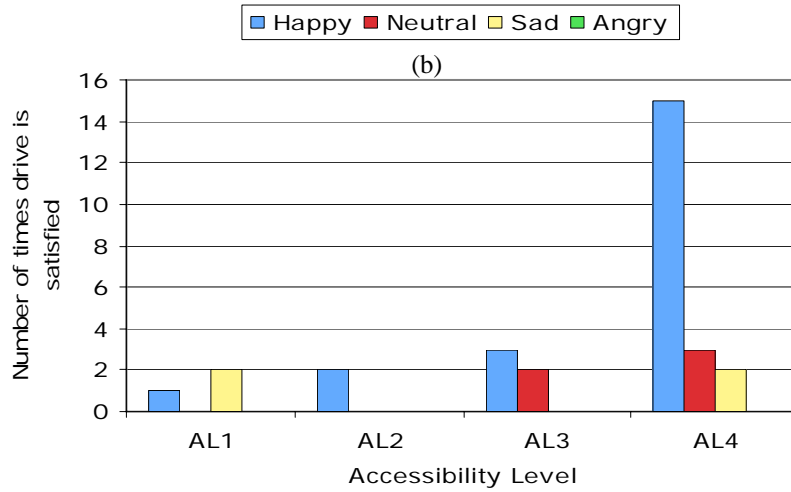


Figure B-6: Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

Figure B-7 – Figure B-9 show experimental results for Experiment #5 with the positive influence factor and the state reliability function for the robot emotion model. 30 participants were utilized in the experiment.

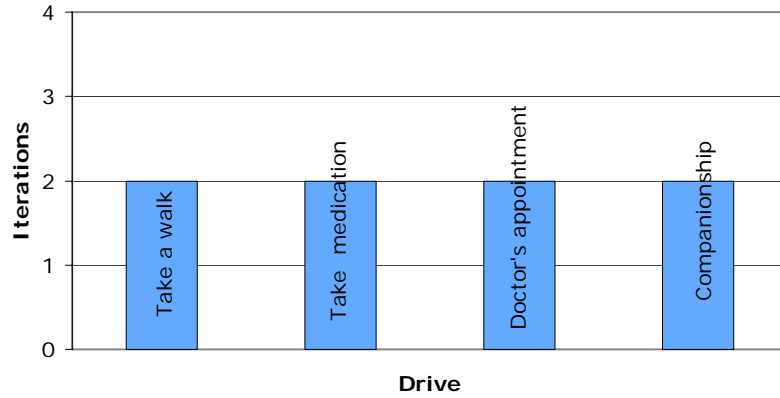


Figure B-7: Average number of iterations needed to satisfy all four drives.

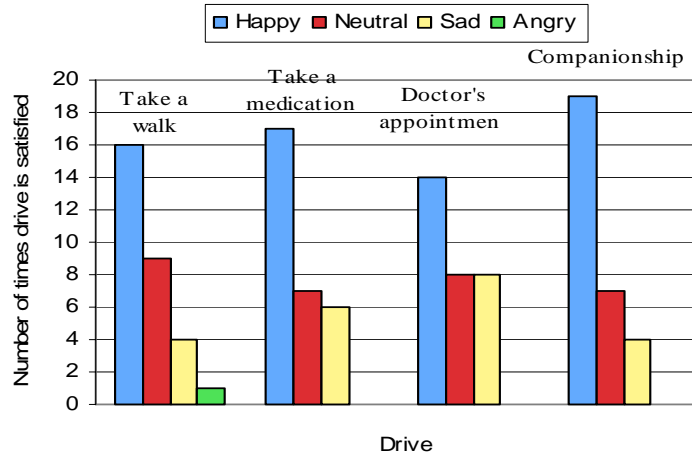
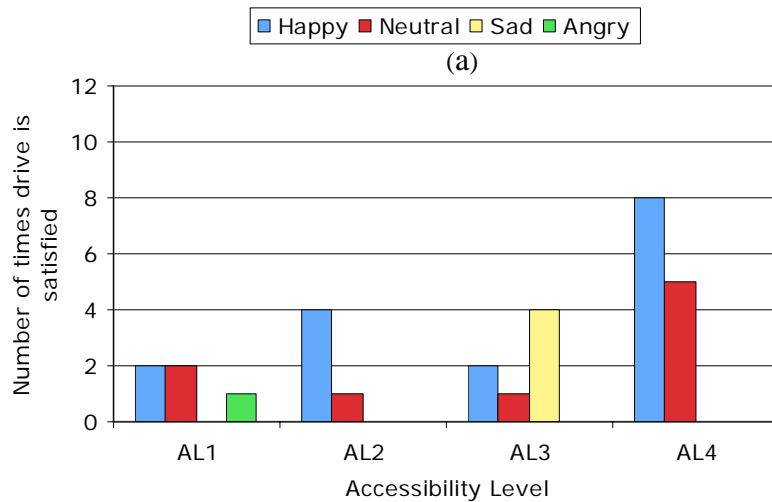


Figure B-8: Emotional states of the robot during drive satisfaction.



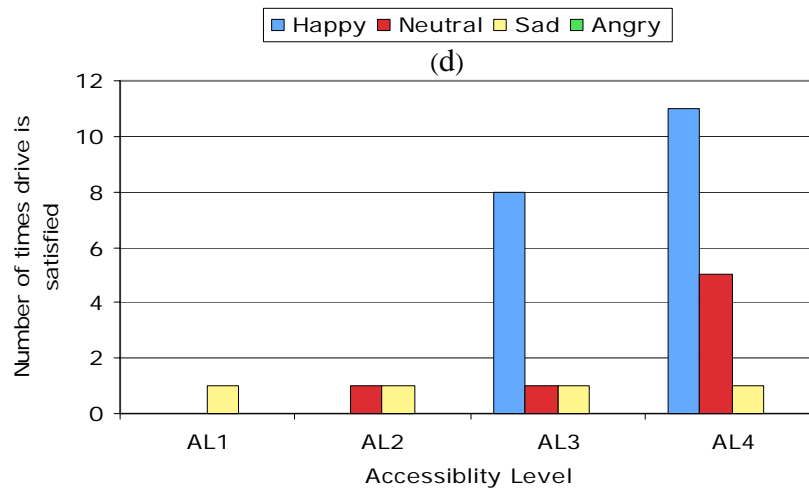
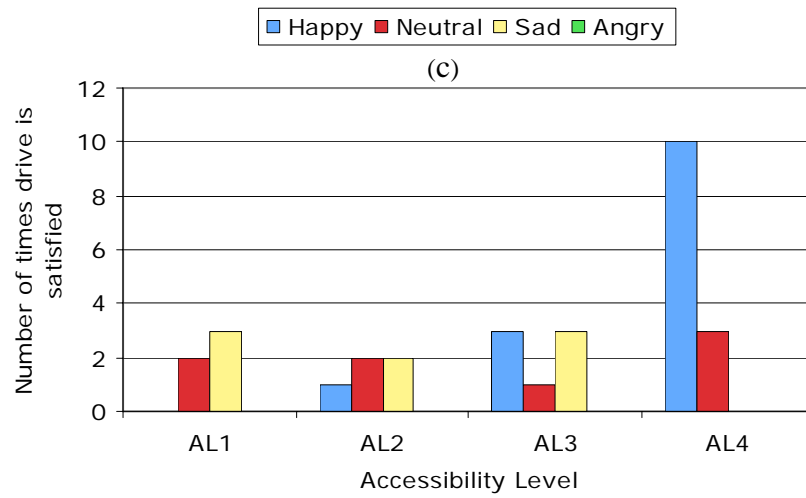
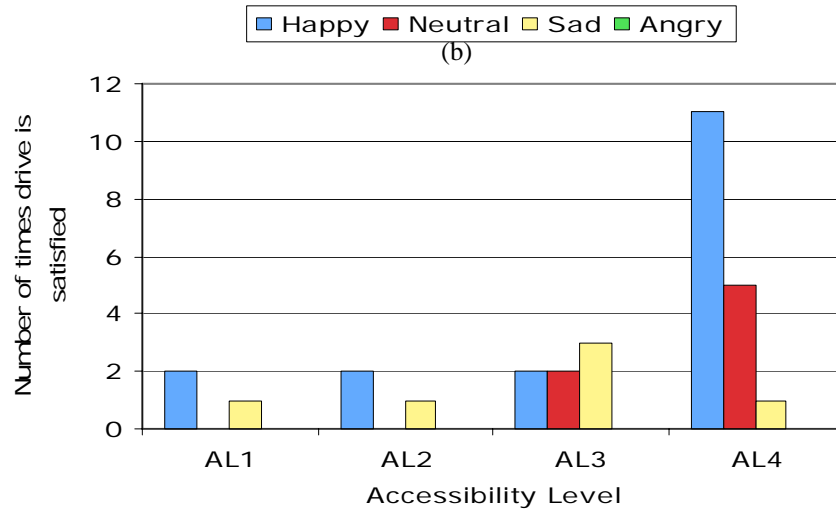


Figure B-9: Frequency of the accessibility levels of the agent when each drive was satisfied: (a) Drive 1, (b) Drive 2, (c) Drive 3, and (d) Drive 4.

Appendix C

A simple example is presented herein to show how the accessibility level of the agent can be determined.

Assume that the “degree of the mood” of the agent is defined by a series of integers:

$$x_H = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}, \quad (C-1)$$

The integers in Equation C-1 representing the “degree of mood” are correlated with the four different accessibility levels via the following probability distribution determined for the agent angry mood state:

$$\begin{aligned} P(x_H = x_{H1}) &= P(C_1) = 0.4 \\ P(x_H = x_{H2}) &= P(C_2) = 0.3 \\ P(x_H = x_{H3}) &= P(C_3) = 0.2 \\ P(x_H = x_{H4}) &= P(C_4) = 0.1 \end{aligned} \quad (C-2)$$

where $C_1=1, 2, 3, 4$, $C_2=5, 6, 7$, $C_3=8, 9$, and $C_4=10$

During interaction between the robot and agent, the accessibility levels of agent are determined based on Equations C-1 and C-2.

The following procedure is implemented:

Step 1: The integer representation (1-10) of the agent’s “degree of mood” is randomly determined using a random number generator..

Step 2: The corresponding accessibility level, as defined by the given probability distribution, which associates with the integer is chosen to be the agent’s accessibility level.

For example, if the integer 5 is picked when the agent is in an angry mood then the agent’s accessibility level will be level II, if integer 10 is pick, the agent’s accessibility level will be level I, etc. A similar procedure is utilized when the agent is in a happy mood.

Moreover, a more precise agent’s mood can be designed by increasing the number of integers for “degree of the mood” (i.e.1-20, or 1-30).

For example:

$$x_H = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20\}, \quad (C-3)$$

$$\begin{aligned} P(x_H = x_{H1}) &= P(C_1) = 0.55 \\ P(x_H = x_{H2}) &= P(C_2) = 0.25 \\ P(x_H = x_{H3}) &= P(C_3) = 0.15 \\ P(x_H = x_{H4}) &= P(C_4) = 0.05 \end{aligned} \quad (C-4)$$

Where $C_1=1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11$, $C_2=12, 13, 14, 15, 16$, $C_3=17, 18, 19$, and $C_4=20$.