Toward Building A Social Robot With An Emotion-based Internal Control

Andreas Marpaung
University of Central Florida

TOWARD BUILDING A SOCIAL ROBOT WITH AN EMOTION-BASED INTERNAL CONTROL AND EXTERNAL COMMUNICATION TO ENHANCE HUMAN-ROBOT INTERACTION

by

ANDREAS HENDRO MARPAUNG
B.S. University of Central Florida, 2002

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the School of Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2004

ABSTRACT

In this thesis, we aim at modeling some aspects of the functional role of emotions on an autonomous embodied agent.

We begin by describing our robotic prototype, Cherry—a robot with the task of being a tour guide and an office assistant for the Computer Science Department at the University of Central Florida. Cherry did not have a formal emotion representation of internal states, but did have the ability to express emotions through her multimodal interface. The thesis presents the results of a survey we performed via our social informatics approach where we found that: (1) the idea of having emotions in a robot was warmly accepted by Cherry's users, and (2) the intended users were pleased with our initial interface design and functionalities. Guided by these results, we transferred our previous code to a human-height and more robust robot—Petra, the PeopleBot™—where we began to build a formal emotion mechanism and representation for internal states to correspond to the external expressions of Cherry's interface. We describe our overall three-layered architecture, and propose the design of the sensory motor level (the first layer of the three-layered architecture) inspired by the Multilevel Process Theory of Emotion on one hand, and hybrid robotic architecture on the other hand. The sensory-motor level receives and processes incoming stimuli with fuzzy logic and produces emotion-like states without any further willful planning or learning. We will discuss how Petra has been equipped with sonar and vision for obstacle avoidance as well as vision for face recognition, which are used when she roams around the hallway to engage in social interactions with humans.

We hope that the sensory motor level in Petra could serve as a foundation for further works in modeling the three-layered architecture of the Emotion State Generator.

I would like to dedicate this thesis to my parents—Simon Marpaung, M.D. and Helena br. Simarmata—for their unconditional love, care, and support both physical and spiritual, and to my brothers and sister—Daniel Petrus Marpaung, M.D., Febri Ronald Marpaung, B.S., and Monika Magdalena br. Marpaung—for their support and love. Without them, I would not be able to be who I am right now.

ACKNOWLEDGMENTS

I would like to acknowledge the following people, who have all been instrumental in the completion of this thesis, and my academic, social, and personal development.

Dr. Christine Lisetti, for her patience, support, guidance and motivation from the minute I took her Affective Computing (AC) class to the completion of this thesis. My interest in the fields of robotic agents and emotion modeling started when I developed and worked with Cherry, as part of my project in the class and they have become my research interests since then.

My other committee members—Dr. Charles Hughes, Dr. David Workman, and Dr. Mason Cash—for their guidance and feedback in making this thesis the best.

My colleagues in the Affective Social Computing Laboratory at the University of Central Florida for their friendship and support.

Eric P. Leger for his tremendous help in coding the navigation and obstacle avoidance system.

Juan Martinez from the University Writing Center at the University of Central Florida for his tremendous help in reviewing and giving feedback.

My wonderful brothers and sisters at City Blessing Ministries, especially Rev. Fredy Liwang, for their love as well as spiritual supports. Thank you very much for showing me what meaning of life is really about.

Above all, I also like to acknowledge and thank God, for choosing me to be one of His sons. Without Him, I cannot live. I depend on Him for saving my soul, giving me abundant blessings, and comforting me with the warmth of His gentle hands every second I need Him.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A: Arousal tag

AI: Artificial Intelligence

AAAI: American Association of Artificial Intelligence

ARIA: Activmedia Robotics Interface for Application

ASCL: Affective Social Computing Laboratory

BSG: Behavior State Generator

CSB: Computer Science Building

EPC: Emotion-like Parameters Change

ESG: Emotion State Generator

FACS: Facial Action Coding System

FLAME: Fuzzy Logic Adaptive Model of Emotions

GRACE: Graduate Student Attending Conference

GSR: Galvanic Skin Response

GUI: Graphical User Interface

HCI: Human Computer Interaction

HRI: Human Robot Interaction

PC: Personal Computer

PDA: Personal Digital Assistant

PETEEI: A Pet with Evolving Emotional Intelligence

S: Stance tag

SECs: Stimulus Evaluation Checks

SM: Somatic Marker

TSK: Takagi Sugeno Kang

USB: Universal Serial Bus

UCF: University of Central Florida

V: Valence tag

CHAPTER 1: INTRODUCTION

> *"The question is not whether intelligent machines can have any emotion, but whether machines can be intelligent without any emotions. I suspect that once we give machines the ability to alter their own abilities, we'll have to provide them with all sorts of complex checks and balances."*
> -- Marvin Minsky (1986)

Although most people realize that emotions play important roles in our lives, the study of emotions, especially in the field of psychology, just started several decades ago. Along with the growing interests in psychology, in the field of computer science, the study of emotions just started recently, especially when Rosalind W. Picard introduced the field of Affective Computing. Picard (1997) defines Affective Computing as "computing that relates to, arises from, or deliberately influences emotions." The domains of this field are very wide, e.g., implementing and expressing emotions in computers and intelligent agents to enhance Human-Computer Interaction (HCI) (Picard, 2001); giving machines the capabilities to recognize emotions by measuring Galvanic Skin Response (GSR), heartbeat and other indicators for future cars (Nasoz et al, 2002) and for telemedicine (Lisetti et al, 2003); or recognizing frustration that leads to learning ability reduction (Lewis and William, 1989).

We usually think of emotions as being nonrational. When we are faced with the fact that someone is an emotional person, we most often carry along a negative connotation toward that person. Some people may question the researchers who want to model emotions in "the emotional agents." Why would computers have emotions? Do we want to have future robots cry because they are depressed with their jobs and stop doing their tasks? Do we want to have our smart houses and appliances angry at us? Or in the worst case, do we want the story in *2001: A Space Odyssey* to happen to us? In this story, HAL, the supercomputer in the space craft that has the capabilities of planning, artificial intelligence, lip-reading, speech recognition and synthesis, commonsense reasoning, recognizing and displaying emotions, and interacting with human naturally, gets angry to the astronauts and at the end tries to kill them. (Stork, 1997). We, as humans, do not want these things happen neither to our future generations nor to ourselves. And we, as the creators, have the full power to limit their emotional states to certain degrees where they still can help us out without jeopardizing ourselves and still behave within some acceptable boundaries.[1]

So, why do we want to think of emotions with computing artifacts?

Computing and Emotion Recognition: Imagine a scenario of an office setting where an employee is working with a slow desktop. After a long and tiring day, the worker is pressured to complete a task before the end of the day. Since his computer does not have faster processors, the software runs really slow. At the beginning, this person may be patient enough to interact with this computer. But the longer the interaction is, the lower his patience level may be, which involves emotions—psychic and physiological changes—which can sometimes (as anger or fear) prepare the body for vigorous actions.[2] If the worker has a meek personality, these actions can be safe, i.e., use other faster computers, but an aggressive personality actions can lead to harmful and damageable ones. To prevent the actions, it would be ideal *if computers can detect the user's*

---

[1] Many questions arise on these acceptable boundaries, which will not be discussed in this thesis.

*emotion*al state somehow, and apply protective measures before any damages may occur. For example, the computer should be able to detect emotion-related physiological changes, i.e., facial expressions, heartbeat, GSR, etc and when it detects any changes, it can suggest that users take a short break to release stress. By doing so, damage can be avoided to computers and any other appliances as well as to humans.

Emotion and Rational Decision-making: It also has been shown that emotion plays an important role in producing rational behavior and decision-making and vice versa (Damasio, 1994). In his research, Damasio found that his patients who had frontal-lobe disorders, which affected a key part of the cortex that communicates with the limbic emotion-related system, always made repetitive disastrous decisions. Although his patients appeared to have normal levels of intelligence (scoring average or even above average on a variety of cognitive tests), when facing real life, doing business for an example, they would make disastrous decisions, although they, previously, had lost a lot of money due to lack of management skills. These behaviors were repetitively performed until all of their capital and collateral was gone. Because of their failures to act rationally, they usually ended up unemployed and living lonely lives. From this finding, it is shown that emotion and rational decision-making are connected. In a simple example, a rational person may associate making a lot of profits with a happy emotion, but on the other hand losing money with a frustration. In the future, if the same situation occurs, the same person will most likely choose the decision associated with the happiness over the

---

[2] Merriam Webster Dictionary (Merriam, 2004) defines emotion as “the affective aspect of consciousness or a psychic and physical reaction (as anger or fear) subjectively experienced as strong feeling and physiologically involving changes that prepare the body for immediate vigorous action.”

frustration one. In this project, we associate the rational decision-making made in an office environment, e.g., avoiding walls, obstacles, etc., to a certain emotion.

Other findings suggest that a little change in emotional state can significantly impact creativity, problem solving, and a willingness to lend a hand to others (Isen et al, 1987; Isen, 2000). With emotions, we can enhance both Human-Computer Interaction (HCI) by designing machine with emotions where the users can be interacting with less stress (Picard, 2001) and Human-Robot Interaction (HRI) by creating interfaces in robots, which integrates domain-specific information and anthropomorphic agents with emotion capabilities in terms of expressiveness and internal states (Lisetti, Brown, Alvarez, and Marpaung, 2004).

This thesis is a research report, which proposes a model of emotion that has been partially implemented on an autonomous robot and which is organized as follows:

- *Chapter 2 – Related Work:* describes the state-of-the-art in robotics' architectures and in modeling emotion for intelligent agents.
- *Chapter 3 – Our Approach:* describes our proposed model of the sensory motor level, based on the *Multilevel Process Theory of Emotion* to model appropriate emotions compatible with the Hybrid Reactive/Deliberative architecture.
- *Chapter 4 – Autonomous Robot Implementation:* describes (1) the creation of Cherry and Lola[3], the first phase to create the prototypes, and (2) the implementation of a PeopleBot™ (ActivMedia, 2002), the continuation of our first phase effort to model

[3] Names and gender (she) are used to personify the robot so the robot is more human-like instead of machine-like to enhance our Human-Robot Interaction goal.

emotion, integrating an avatar and text-to-speech engine (Haptek, 2002), face recognition (Identix, 2002) and navigation and vision systems into a user friendly Graphical User Interface (GUI).

- *Chapter 5 – Experimental Results:* exposes and explains the collected data in the *UCF Affective Social Computing Laboratory* and the robot's intended world, and an analysis of the results.
- *Chapter 6 – Future Work and Conclusion:* discusses suggested future works.

1.1. Research Question

The research question that this thesis will address is thus: How can the stimuli from outside world be processed at the sensory motor level to have effects on the emotional state of an intelligent agent?

1.2. Contributions

- To identify latest multimedia technologies that are necessary for social interaction (e.g. face recognition, speech, facial displays, emotional expressions, knowledge of people's status and etiquette rules);

- To integrate these multimedia technologies into a multimodal interface that can help us to enhance Human-Robot Interaction (HRI) from the social interaction perspective; i.e., for the social robot operated in an office environment, we use a pleasant anthropomorphic female agent;
- To evaluate the user's acceptance of such an anthropomorphic interface in our specific context;
- To construct a robotic multimedia platform (e.g. sonars, laser, vision) for an office robot with various specific tasks (e.g. tour guide, master-gopher) that can be tested "in vivo" with real users;
- To enhance the simple multimedia robotic platform by conceptualizing a three-layered architecture based on emotion theory for control of internal states;
- To design and implement one of the layers of the three-layered architecture using formal representations and scripts of emotion-like states, both at the internal level of progress toward goals, and at the external expression level of facial expressions and text;
- To suggest ways in which the other two layers of the architecture can be linked with the first for future research.

CHAPTER 2: RELATED WORK

The growth of Artificial Intelligence, especially the field of Affective Computing, has made many researchers become interested in studying the effects of emotions in social interactions between the users and the social agents (Lisetti, Brown, Alvarez, and Marpaung, 2004), interactions between or among agents (Murphy, 2002; Michaud, 2001; Michaud et al 2000, 2001a, 2001c), and social exchange and learning in agents (Breazeal and Scassellati, 2000; Breazeal, 2003). To study these effects, several models of emotions have been proposed and implemented in various test-beds, i.e., simulations and robots with different architectures. This section is dedicated to explaining: (1) the state-of-the-art of the robotic architectures, and (2) several research projects that model the emotions from different perspectives for different goals and purposes.

2.1. Robotic Architectures

There are many controversies among the researchers in deciding the architecture most suitable for many different applications. This subsection is dedicated to explaining the existing main architectures in the robotics domains–hierarchical, reactive and hybrid deliberative/reactive.

2.1.1. Hierarchical Robot Architecture

The hierarchical architecture is the oldest method of organizing intelligence in robotics. This paradigm is defined by the relationship among three primitives—sense, plan and act—which are executed in the sequential order as shown in Figure 1. Initially, the robot *senses* the world, and then *plans* the actions based on the sensed information. Once the plans are laid out, the robot can *act* based on the formulated plans.

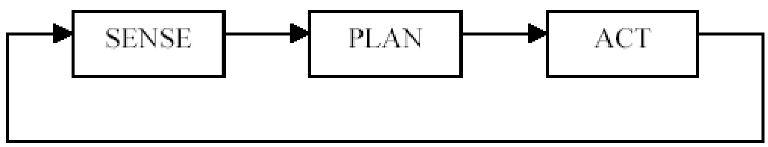


Figure 1: Hierarchical Paradigm (adapted from Murphy, 2000)

The organized execution between sensing, planning and acting made this paradigm popular before 1990. The first AI mobile robot that uses this paradigm is Shakey the Robot, a robot that needs a generalized algorithm for planning in order to accomplish goals (Nilsson, 1984). Shakey specifically uses the *strips*[4] as part of the General Problem Solver method (Newell and Simon, 1972) with means-ends analysis approach.

---

[4] Strips uses the means-ends analysis approach, where if the robot cannot reach the goal in one movement, it picks an action, which will reduce the difference between what state it was in now and the goal state.

This architecture introduces two problems: the *closed world assumption* and *frame problem* (Murphy, 2000). In the *closed world assumption*, no surprises can be allowed, thus making the identification of the combination of all possible cases necessary. The *frame problem* arises, on the other hand, when another possible medium (or environment), i.e., additional obstacles or rooms to navigate are added into the current medium. As a simple example, the number of possible cases for a robot to consider in order to move around and determine its pathway is higher when a third room is introduced to a current two-room medium (or environment).

As the field of robotic research attempts to deal with uncertainty and abilities to react to unsuspected outcomes in order to increase robot autonomy, the need to have an open world assumption—allowing for changes to occur during execution of planned actions—is increasing. The open world assumption is thus more realistic than the closed world assumption of the hierarchical paradigm. In addition, generating all possible cases can be computationally slow and hence very cumbersome.

2.1.2. Reactive Robot Architecture

A new robotic paradigm was therefore introduced to address some of the limitations of the hierarchical architecture–namely, the reactive architecture. In this paradigm, the plan component is literally put aside. The other two components—sense and act—are tightly coupled and the interactions between these two primitives can give specific behaviors. The reactive architecture is shown in Figure 2.

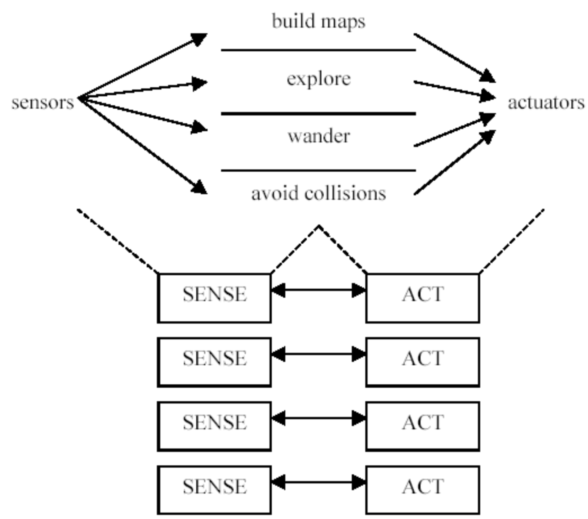


Figure 2: Reactive Paradigm (adapted from Murphy, 2000)

Programming with this architecture is often referred to as “programming by behavior”, since the main component of the implementation is behavior. Increasing the number of simple behaviors, the robots can become more intelligent by choosing the best solutions among them based on the stimuli accepted that are dynamically changing.

The most influential of purely Reactive systems is the *Subsumption Architecture* introduced by Rodney Brooks in some shoebox-sized insect-like robots with six legs and antennae (Maes and Brooks, 1990). In the subsumption architecture, the network of sensing and acting modules are grouped into layers of competence. Lower layer acts as the basic functions such as obstacle avoidance, and the higher layer performs more goal-directed actions such as mapping. Tasks in this paradigm are accomplished by activating the appropriate layer, based on the agent's current conditions and needs, which can then activate the lower layers below it.

In the subsumption paradigm, the closed world assumption and the frame problems are solved. Indeed, because behaviors do not monitor the changes in the environment but rather simply responds to whatever stimulus is in the environment, using this paradigm the world does not need to be assumed closed. Moreover, the frame problems are fully eliminated because the need to model the world is also eliminated. However, only relatively simple behaviors can be produced.

Although the reactive architecture is more popular than the hierarchical one, it is limited to applications which can be dealt with reflexive behaviors such as knee-jerk (Biological example) and cannot be used for applications that require planning, reasoning about resource allocation such as obstacle avoidance system.

2.1.3. Hybrid Deliberative/Reactive Robot Architecture

The controversies between the lack of planning and reasoning in the reactive architecture on the one hand, and the closed world assumption and frame problems in the hierarchical architecture on the other hand, brought about in the late 1990 (Arkin, 1998) the conceptualization of another type of architecture called *Hybrid Deliberative/Reactive Architecture.* The organization of this architecture can be described as *plan* and then *sense-act* as shown in Figure

3. In this architecture, a robot plans its possible actions and instantiates a set of behaviors to execute that plan. The behaviors in the plan can be executed until the plan is completed. The planner then recursively generates a new set of behaviors and those behaviors are executed.

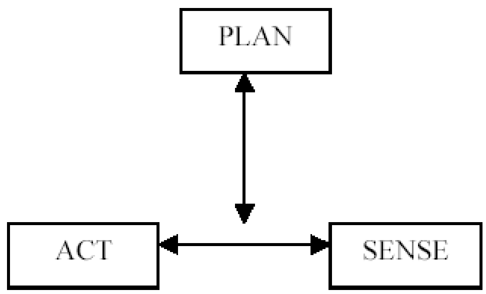


Figure 3: Hybrid Architecture (adapted from Murphy, 2000)

The hybrid architecture hence combines both reactive and deliberative components. The reactive portion uses local and behavior-specific representations while the deliberative one uses global world representations. Behaviors in the reactive architecture are different in nature from the ones in the reactive portion of the hybrid architecture. In the reactive paradigm, the behavior is purely reflexive behaviors whereas in the hybrid paradigm, the behavior includes both reflexive and learned behaviors. Although when this paradigm was introduced, many researchers thought this architecture as a theoretical one, but now there are many robotic implementations actually use this architecture (Murphy et al, 2002; Michaud, 2001; Lisetti et al, 2004; Breazeal and Scassellati, 2003).

2.2. Emotion Modeling Projects

In this section, several research projects that have attempted to begin modeling emotion on a variety of platforms (single robot, autonomous cooperating robots, software agents) are discussed.

Table 1 (at the end of this section) shows the differences among the emotional agents' projects, explained in details below, in terms of hardware, architecture, functionalities, and modes of interactions.

2.2.1. Kismet

A computational model of emotion has been developed in *Kismet*, a robotic face with some degree of intelligence, built at the Media Laboratory at the Massachusetts Institute of Technology (Breazeal, 2003). Kismet, shown in Figure 4, interacts with its caretakers by perceiving a variety of inputs from its visual and auditory channels and gives feedback through its gaze direction, facial expressions (Figure 5), body posture, and vocal babbles.

To interact with users, Kismet has five different subsystems: perception, motivation, attention, behavior, and motor (Breazeal, 2000). The *perception system* handles Kismet's vision, which is capable of perceiving motion and can recognize faces in some simple manner. It extracts sensor-based features from the outside world and encapsulates these features into percepts of social stimuli such as faces, voices, as well as salient non-social stimuli, such as brightly colored objects, loud noises, large motion which can influence behavior, motivation, and motor processes.

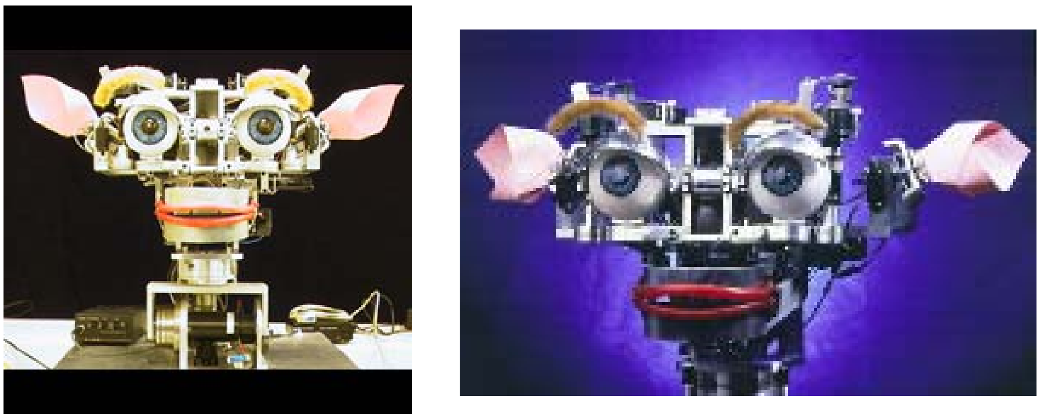


Figure 4: Kismet (adapted from Breazeal, 2002)

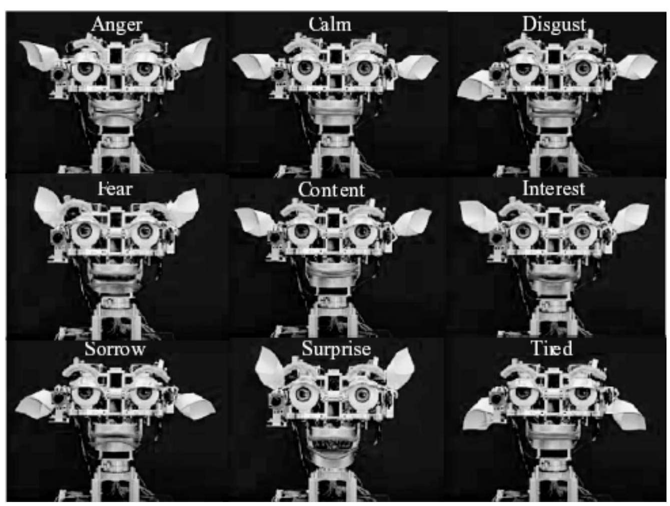


Figure 5: Kismet's facial expressions (adapted from Breazeal, 2003)

The *motivation system* consists of two inter-related subsystems: *drives* and *emotions*. The *drives* represent the robot's needs and are modeled as simple homeostatic regulation mechanisms. The homeostatic regulation mechanisms are the processes by which the critical parameters (e.g. temperature, energy level, amount of fluids) are kept within a bounded range in order to avoid damage. Kismet has three basic drives: (a) the social drive or the need to socialize; (b) the stimulation drive or the need to respond to stimuli generated either externally by the environment or internally through self-play; and (c) the fatigue drive or the need to take a rest after socializing for a certain period of time by shutting itself off from the external world. When the needs are met, the intensity level of each drive falls within a desired range. On the other hand, when its needs are not met, the drive's intensity either increases or decreases, depending on its internal state and the inputs accepted by the perception system. In short, these drives represent the robot's own agenda and do play significant role in activating certain behavior at certain time. The *emotions*, on the other hand, show its internal states that can be displayed through its facial expression, body posture, and tone or vocal babbles. Kismet can also display several other responses, such as interest, calm, and boredom that correspond to the inputs that have high arousal values

The *attention system* receives low-level visual percepts from the perception system. This system is able to pick out low-level perceptual stimuli, relevant at that time, to direct its attention to that related stimuli immediately. Sudden appearance, sudden change, and inherent saliency are several ways to get its attention. The *behavior* system has several components associated with each drive in the motivation system. They are: *socialize* acts to move the social drive, *play* acts to move the stimulation drive, and *sleep* acts to satiate the fatigue drive. This system activates certain behavior(s) while maintaining relevancy and persistence. And finally, the *motor* system deals with the robot's motor skills and expressions so the robot can carry out its course of action based on its internal states or emotive expressions. This motor system is built by four other subsystems: *the motor skills system* that deals with motor functionalities, *the facial animation system* that deals with facial muscles' movements, and expressive facial displays, *the expressive vocalization system* that adjusts the tone and lips position based on the triggered emotion, and *the oculo-motor system* that moves its eyes to the target chosen by the attention system.

When Kismet senses something from the outside world, the inputs are accepted, filtered by several extractors and encoded by releaser processes. In the encoding processes, releasers are evaluated and the results are set to their activation level. If the results are above certain thresholds, then the inputs are passed to their corresponding behavior processes in the behavior system and to the affective appraisal if they can influence its emotion system. In the affective appraisal stage, the inputs are tagged with somatic marker (SM) tags: an arousal tag (A) to specify how arousing the inputs are, a valence tag (V) to decide how (un) favorable the inputs are, and a stance tag (S) to decide how approachable the percepts are. After tagging the inputs, they are passed to the emotion elicitor stage that will be continued with the calculation of the activation level of each emotion process. This activation level produces the end result of emotion, which is also mapped in the affect space (arousal-valence-stance) that is then sent to the motor system.

Kismet and Petra are similar in that they both are able to display their emotion-like states through their facial expressions. They are also developed in order to learn social interactions between the users[5] and the robots. Currently, Kismet has six basic emotions, formulated by Ekman's Facial Action Coding Systems (FACS)[6] (Ekman and Friesen, 1978)—anger, disgust, fear, joy, sorrow, and surprise. Petra also has five basic emotions, also based on Ekman's FACS—happiness, surprise, fear, sad, and anger.

Despite these similarities, their main difference is their maneuverability. Kismet is designed to be stationary while Petra can explore her environment. With this advantage, Petra can be exposed to a larger world in which the robot is able to navigate, learn, and label social cues more flexible in contrast to Kismet in its smaller and limited world.

2.2.2. Leguin & Butler

Butler and Leguin, shown in Figure 6, are two Cooperating Heterogeneous Mobile Robots that were developed at the University of South Florida, Tampa (Murphy et al, 2002). They won the Nils Nilsson Award for Integrating AI Technologies at the AAAI 2000 Mobile Robot Competition's *Hors D'Oeuvres, Anyone?* At this competition, the waiter robot, Butler,

---

[5] Kismet's users are the caregivers and Petra's users are the visitors, faculty, staff, and students of UCF.

[6] FACS is a system that determined the contraction of each facial muscle (singly and in combination with other muscles) changes the appearance of the face.

had to serve the attendees with hors d'oeuvres while the refiller robot, Leguin, waited for Butler's request for the refill.

Butler has four strategic behaviors: serve, exchange, intercept, and goBack; two monitors: treat-monitor and tray watch; and four emotions: happy, confident, concerned and frustrated. Leguin has four strategic behaviors: refill, exchange, wait, and goHome; no monitors; and three emotions: happy, confident, and concerned.

Figure 6: Butler (left) and Leguin (right) (adapted from Murphy et al, 2002)

At the instantiation, Butler begins with the serve behavior with happy as her[7] initial emotion. During this behavior, she finds faces and serves the attendees she encounters. Also when the serve behavior is active, the treat-monitor is also active. When the treats reach a certain threshold amount while serve behavior is active, her emotional state changes to a more negative one—confident. When the emotion state changes to the more negative one, she sends the message to Leguin to ask her for the refill. If Leguin does not respond to the call due to communication loss, for example, her emotional state becomes concerned and makes her issue the hurry command to Leguin, and finally, she becomes frustrated, which triggers the activation of intercept behaviors to locate Leguin and get the refill by herself. When Leguin's position is found, she finds her way to Leguin. Once they are close to each other, the exchange behavior is triggered and her serve behavior will be re-activated and her emotion is set back to happy.

Leguin, the refiller robot begins with the wait behavior and happy as her initial emotion. When she receives a request from Butler, she initiates her refill behavior and runs at a safe speed. But if she receives a hurry command, she runs at the maximum safe speed. When they are within 1.5 meters, the exchange behavior occurs and she will perform the goHome behavior at a safe speed.

In this project, emotions produced by the Emotion State Generator (ESG) are modeled in order to avoid any deadlocks. This ESG is then linked to the Behavior State Generator (BSG) in which the robot can choose appropriate behavior to a certain emotion from the scripted schemata. Without these emotions, there will be a circular-dependency between the robots if

---

[7] Murphy et al's (2002) and our approach are the same in that we use the names and female gender.

unwanted situations happen, e.g., communication loss where Butler keeps sending messages to and waiting for Leguin while Leguin keeps waiting for Butler's message.

Our approach is similar to Murphy's et al's (2002) in that we use the multilevel hierarchy of emotions[8] where emotions both modify active behaviors at the sensory-motor level and change the set of active behaviors at the schematic level. Despite the similarity, our approach is different from Murphy's et al's in that our robot is able to display emotions through the anthropomorphic face.

2.2.3. Fuzzy Logic Adaptive Model of Emotions (FLAME)

El-Nasr has developed FLAME, a new computational model of emotions in an animal simulation—PETEEI (A Pet with Evolving Emotional Intelligence) (El-Nasr, 2000). PETEEI, shown in Figure 7 and written in Java, models a dog in a graphical interface with five major scenes: garden, bedroom, kitchen, wardrobe, and living room; several feedbacks: barking, growling, sniffing, etc.; and several actions: looking, running, jumping, playing, etc. Under these situations, users can stimulate various behaviors and situations by clicking on the buttons: *walk to different scenes* to move to another scene, *object manipulation* to take an object from the scene, *talk aloud* to talk to the objects, *opening and closing doors* to open and close objects, *look at* to examine the objects, and *touch and hit* to touch and hit the objects.

---

[8] Multilevel hierarchy of emotions is explained further in the "Approach" section.

Figure 7: PETEEI (adapted from El-Nasr, 2000)

FLAME is implemented with fuzzy logic where the agent is capable of mapping events and observations to emotional states and representing emotions based on the calculated intensity. This system also includes an inductive learning algorithm that allows the agent to adapt its response dynamically. The agent receives external events through click-on buttons that are passed to both learning and emotional components as perceptions. Along with the outcomes from the learning component and knowledge of event-goal and expectations, the emotional component produces the emotional behavior that soon becomes the input for the decision-making component to choose appropriate action in response to the previous event. In modeling emotions, the simulation determines which goals are affected and the degree of impact, and calculates the desirability level and the importance of the goals involved. These steps produce the desirability, measured by several fuzzy logic sets, that is passed to the appraisal process to change the emotional state of the agent.

Although both FLAME and Petra use fuzzy logic in modeling emotions, the models are implemented in different test beds—FLAME is in a simulation and Petra is in a social robot. In its limited world, FLAME can only accept the inputs from its click-on buttons without measuring the degree of inputs. Mainly, El-Nasr's research is directed toward Human-Computer Interaction (HCI) and computer simulation. In contrast, ours is directed toward Human-Robot Interaction (HRI).

2.2.4. Cathexis

Another model of emotion is Cathexis (concentration of emotional energy on an object or idea) proposed by Velasquez (1996) to simulate emotions in autonomous agents. The emotions are modeled as a network of special emotional systems comparable to Minsky's (1986) "proto-specialist" agents. Each proto-specialist that represents both cognitive and non-cognitive emotion activation systems is built by four different groups: *neural,* which covers the effects of neurotransmitters, brain temperature, and other neuroactive agents; *sensorimotor,* which covers sensorimotor processes; *motivational,* which covers drives, emotions, and pain regulation; and *cognitive,* which covers the cognitions that activate emotion. Beside those four groups, each proto-specialist also has $\psi(\ )$, the emotion decay function that controls the emotions' lifetimes;

two threshold values: α that controls the activation of the emotion and ω that controls the saturation level of the emotion, and finally I that controls the emotions' intensities.

The intensity of emotion is calculated by summing up the previous level of arousal, the contributions of each emotion elicitors, and the interactions of inhibitory and excitatory inputs that is then sent to the function that constrains the intensity of an emotion between 0 and its saturation value ω.

Butler and Leguin are similar to a system based on a Cathexis model in that their emotion systems are linked to the behavior system. From this behavior system, the agent can choose appropriate behavior that depends on the internal state. This behavior system has an expressive or a motor component that contemplates prototypical facial expression, body posture, and vocal expression and an experiential component that is built from the motivation and action tendency.

This model was implemented in Simon the Toddler, shown in Figure 8 (Velasquez, 1997). It was a synthetic agent in which Simon had five drive proto- specialists–hunger, thirst, temperature regulation, fatigue, and interest–and six emotion proto-specialists—enjoyment or happiness, distress or sadness, fear, anger, disgust, and surprise. Simon accepts stimuli from the changes of different controls set by users, i.e., sliders, icons, and buttons that alter Simon's neurotransmitters' and hormones' levels.

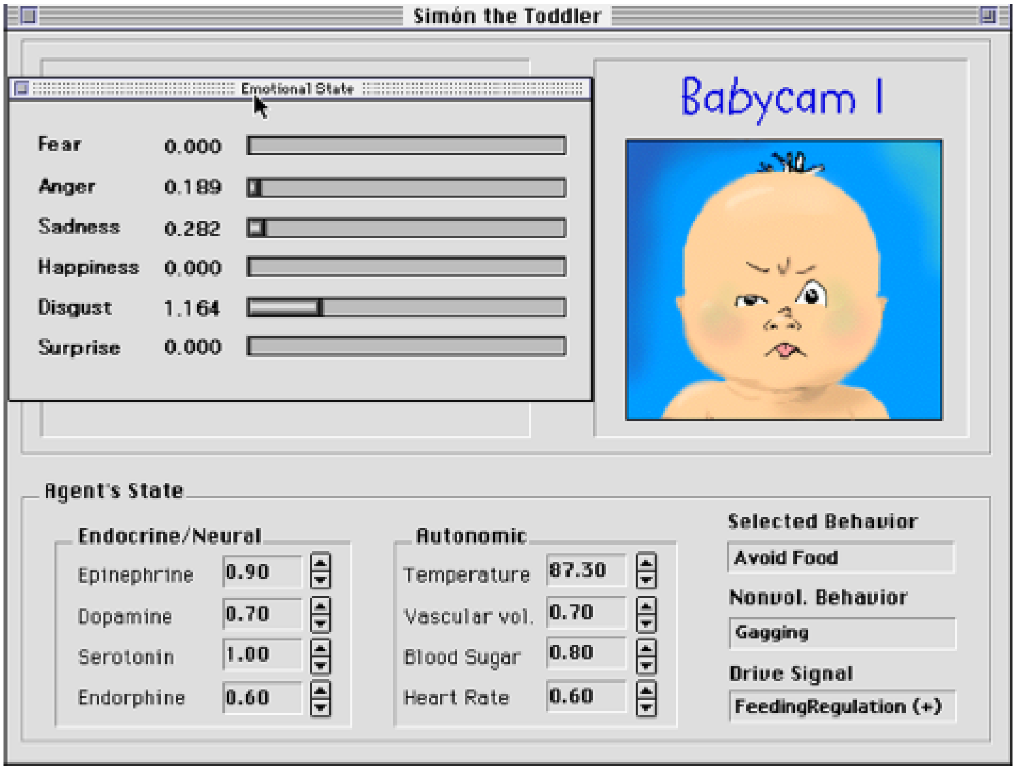


Figure 8: Simon the Toddler's interface (adapted from Velasquez, 1997)

Another model Emotion-based Decision Making is also proposed by Velasquez (1998) as an extension to the Cathexis model. This more complex model is implemented in the Virtual Yuppy, shown in Figure 9, a simulated emotional pet robot. In this model, Velasquez integrates several other systems such as perceptual, drive, and enhanced motor systems.

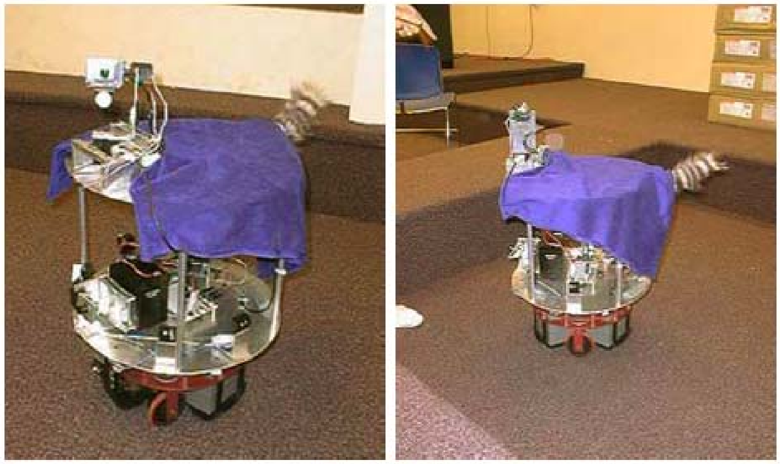


Figure 9: Virtual Yuppy (adapted from Velasquez, 1998)

2.2.5. Emotions in a Group of Autonomous Robots

Not only are artificial emotions being implemented in a single agent, but modeling the emotions in a group of autonomous robots also has been explored (Michaud et al, 2001c; Michaud, 2001). Michaud et al's works are influenced by Maslow's Hierarchy of Needs Theory, which integrates the physiological needs, such as hunger, thirst, breathing, and sleeping, into their model. In this work, the robots have the intelligence to search for and dock into a near-by charging station and to decide on when and for how long they should recharge themselves using the motives[9] and artificial emotions to regulate this process.

---

[9] Motives and goals are interchangeably used in this section.

In the proposed work, the emotion is modeled to monitor the motives, which are associated with the accomplishment of the robots' goals in a certain time frame. For example, when the robots' energy levels are decreased to certain threshold (influenced by battery voltage level, sensing of the charging station, and rational module) or the activation level of *Energize* is greater than 50 %, the *Recharge* behavior is desired. Under this circumstance, the *Energize motive* is activated in order to determine when and for how long the recharging process is needed. When the recharging is not needed, the *Follow-Wall* motive is set to desirable.

This artificial emotion model is programmed in several Pioneers 2 robots that use a hybrid reactive-deliberative architectural methodology, shown in Figure 10a, which is built on top of behavior-producing modules connecting sensory information to commands. These collections of modules select the behaviors dynamically based on the influences that are coming from the *Implicit* conditions from the sensory inputs, the organization of goals, managed by the *Egoistical* module, and the reasoning that comes from the *Rational* module.

In his work, Michaud models four different emotions mapped in a two-dimension bipolar coordinates: joy and anger (positive emotions) and sadness and fear (negative emotions) that can be shown through their interfaces (Figure 10b). Both joy and anger show whether the robots can accomplish the goals. Joy shows its achievements by monitoring the decreasing energy level, while anger, on the other hand, shows them by monitoring the energy level oscillations. In contrast, both sadness and fear show whether the robots have difficulties in achieving the goals. Opposite to joy, sadness does monitor an increase in the energy level. The higher the level, the higher the chances the robot wants to change its strategy. In parallel to sadness and opposite to anger, fear shows its difficulties by monitoring its constant energy level.

(a) (b)

Figure 10: (a) Pioneer 2 robots, (b) Facial Expressions of the robots (adapted from Michaud, 2001b)

Our approach is the same as Michaud et al's in that we use artificial emotions to monitor the agents' accomplishments toward the goals. Despite the similarities, our work is defined in a more formal cognitive model than his previous proposed model by sensing, collecting, measuring, and calculating many different stimuli, from several sensors, that may affect the emotion-like state.

2.2.6. Graduate Student Attending Conference (GRACE)

GRACE, shown in Figure 11, is a B21R Mobile Robot built by RWI that is also used as a social robot (Simmons et al, 2003). It has an expressive face on a 15” flat-panel LCD screen mounted on a panning platform with a large array of sensors that include touch, infrared, and sonar sensors, a microphone, a scanning laser range finder, a stereo camera head on a pan-tilt unit, and a single color camera with pan-tilt zoom capability. GRACE can also speak using a high-quality speech synthesizer, and understand responses using its microphone and speech recognition software.

Figure 11: GRACE (adapted from Simmons et al, 2003)

GRACE is an entry at the 2002 AAAI Robot Challenge where the robot must find the registration booth, register, interact with people, and, with a map handy, it needs to find its way to a certain location to give a technical talk about itself. The main objectives of the challenge are to: (1) demonstrate a high level of intelligence and autonomy for robots to act naturally and be people-oriented in a dynamic environment; (2) integrate the current state-of-the-art equipments to solve this task; and (3) demonstrate and educate the intended audiences about the difficulties and challenges in robotic research.

As part of the challenge, GRACE started at the entrance door and found its way to the elevator by either interacting with people or performing a random walk until a person was encountered. For the interactions, it used: (1) the IBM's VIAVOICE, which converted the speech to text strings that would be parsed by NAUTILUS, Naval Research Laboratory's natural language-understanding system, and then mapped into a series of messages, and (2) a personal digital assistant (PDA), which received the directional gestures input, i.e., turn left and take the elevator to the second floor. Through these two modes, the directions were received and inserted into a queue to be processed, and finally, the robot stood close to the elevators. When it thought that the distance between itself and the elevator was close enough, it needed to find the closest elevator among the three available ones by performing and processing the laser readings. After finding the elevator, it had to wait for an open door and entered the elevator when it was open. With the help of humans or other means, it could reach the destination floor and got out from the elevator when the door was open.

After reaching the destination floor, the robot needed to find the bright pink registration sign through the Swarthmore vision module with the Canon VC-C4 pan-tilt camera. And when the distance between the pink sign and the robot was close enough, the robot had to stand in line. In order to keep acceptable personal space between people in line and the robot, it had to perform a SICK laser scan. It kept maintaining the distance until it reached the end of line known from the laser scanning (assuming that it stood next to the registration desk where it had to find a person and to register for the talk). During the registration process, it received the map that could be used to navigate to the talk room. After completing the registration's procedure, instead of using the given map, it found its way using its own built-in map. After reaching the room, it had to give the presentation about its technologies, which was done by the Northwestern team.

The avatar that GRACE has is generated by computer, and currently, its facial expressions have not been linked yet to its internal state. In contrast to our work with emotion representations, GRACE does not have any mechanism that can generate several different emotions.

Table 1 below shows the summary and categorization of the projects, explained in this related works section, in terms of hardware, architectures, functionalities, and modes of user interactions.

Table 1: Summaries of related works

Project's names	Hardware	Architecture	Functionalities	Modes of User-Interactions
Kismet (Massachusetts Institute of Technology)	Personal Computers (PCs), 4 color CCD Cameras, Microphones	Hybrid Deliberative/ Reactive	Interacts with its caretakers and gives feedback through its gaze direction, facial expression, body posture, and vocal babbles	Cameras, Microphones, Speech Synthesizer, Facial Expressions, Body Postures
Leguin and Butler (University of South Florida)	2 Nomad 200 robots, Dual Hitachi Color Cameras	Hybrid Deliberative/ Reactive	An entry at the AAAI 2000 Mobile Robot Competition's that served the participants with hors d'oeuvres	Laser beams, Speech Synthesizer
FLAME (Texas A& M University)	Personal Computer (PC)	N/A (this is a software agent NOT a robotic agent)	Simulates the model of emotions in PETEEI based on the chosen buttons	Click-on buttons
Cathexis - Simon (Massachusetts Institute of Technology)	Personal Computer (PC)	N/A (this is a software agent NOT a robotic agent)	Simulates emotions based on the cognitive and non-cognitive emotion activation systems selected by the users	Sliders, Icons, Buttons, Facial Expressions

Cathexis – Virtual Yuppy (Massachusetts Institute of Technology)	CCD cameras, Microphones, IR sensors, Air pressure sensor, Pyro sensor	Hybrid Deliberative/Reactive	Finds the pink bone and interacts with the person who holds it, and interacts with the environment, i.e., darkness, blue Styrofoam horses	Cameras and Pyro sensor to detect changes in temperature due to the presence of humans
Autonomous Robots (Universite of Sherbrooke)	Pioneers 2 robots, Cameras	Hybrid Deliberative/ Reactive	Stimulates emotions based on the energy level and motives or goals.	Does not interact with the users
GRACE (Carnegie Mellon University, Metrica, Inc, Naval Research Laboratory, Northwestern University, and Swarthmore College)	B21R Mobile Robot (comes with touch, infrared, sonar, and SICK scanning laser range finder), Microphones, Sony VAIO Picturebook, Metrica TRAClabs Pan-tilt camera, Canon Single Color Camera with Pan-tilt-zoom Capability	Reactive	An entry at the 2002 AAAI Robot Challenge in which the robot had to find its way to the registration desk, register, and give a talk in front of the attendees.	Microphone, PDA, Cameras

CHAPTER THREE: OUR APPROACH

Motivated by Minsky's quotation (1986)

> "The question is not whether intelligent machines can have any emotion, but whether machines can be intelligent without any emotions. I suspect that once we give machines the ability to alter their own abilities, we'll have to provide them with all sorts of complex checks and balances."

and related work in AI robotics, we propose an approach to integrate emotion-like states into robotic platforms within a relatively simple context. As will be explained in the implementation section, our approach has been two-phased: (1) we first develop a simple robotic prototype simulating simple emotions mostly at the level of *external communicative behavior* through a multimodal interface that we suitably designed and evaluated for social interaction and for specific domain-tasks, and (2) we study, design and partially implement a three-layered architecture based on cognitively grounded theories of emotions to simulate internal motivational goal-based activities.

This chapter is dedicated to (1) describing the (simple) *social expertise* we intend our robots to have in terms of external communicative behavior and motivational goal-based activities; and (2) presenting the psychologically grounded theory of emotions that inspire our three-layered architecture design and implementation, and multimodal interface.

3.1. Basic Social Expertise Activities

As our initial approach in modeling emotion, we build Cherry, the red AmigoBot™, whose functionality we design so that she can socially interact with humans on a daily basis in the context of an office suite environment. Cherry has a given set of office-tasks to accomplish, from giving tours of our computer science faculty suite to visitors, to serving beverages to those faculty and staff, and to engaging them in social interaction.

In addition, Cherry is also being designed to have a variety of internal states and external behaviors such as: (1) maintaining and expressing a consistent personality throughout the series of interactions; (2) experiencing different inner emotional-like states in terms of her progress toward her goals; (3) choosing (or not) to express these inner states in an anthropomorphic manner so that humans can intuitively understand them; (4) having an internal representation of her social status as well as the social status of her "bosses"; (5) adapting to the social status of the person she is interacting with by following acceptable social etiquette rules.

To accomplish her tasks, she is equipped with the second floor map, navigational system, face recognition algorithm, a database of images that are integrated in the interface with an avatar, the anthropomorphic face. Each suite in the map, which is displayed in the interface, is associated with its x- and y- positions (in accordance to the elevator) and the faculty or staff member's information that resides in that suite. With the navigational system, Cherry is able to roam around the hallway to complete her task while avoiding any obstacle that she encounters. The face recognition algorithm enables her to take pictures of person encountered, compare the captured images with the existing database (that store facial picture images with their corresponding names and social status), and greet them based on the etiquette rules. For instance, a Full Professor is greeted with more deference than a 'mere' Graduate Student, following some social rules given to her.

In addition, in order to facilitate the social interaction with humans, an anthropomorphic avatar has been created for Cherry. The avatar is present on the laptop/Cherry's user interface and has voice ability so that she can speak to the user in natural language. She explains a variety of facts, from who she is and what her mission is, namely the UCF computer science tour guide, to which professor works in what office, to what that particular professor is researching.

Through this avatar, she is also capable of showing her internal state appropriately that corresponds to a certain facial expression. These inner states—measured in terms of her current relationship with her environment and goals—will need to be integrated with the external behavior for a consistent system (Ortony, 2001). Currently, Cherry can display different facial expressions corresponding to her different inner states:

(1) Happy: Cherry expresses happiness when she is successful in achieving her goal.

(2) Frustration: Cherry expresses frustration when she finds that the office to which she is sent to has its door closed, or the door is open but she cannot recognize the faculty inside the office.

(3) Discouragement: Cherry shows discouragement when, after waiting for a while (a parameter of her patience which can be adjusted), the door remains closed.

(4) Anger: Cherry can also express anger when, after waiting for a long time, the door still remains closed. This option is created in order to test how people may react to her anger

differently: some may want to help her accomplish her goal, whereas others may not want to deal with her at all.

Having given a quick overview of the various activities envisioned for our service robots, we now describe the psychological theories that have inspired our approach.

3.2. Emotion State Generator

The experience of emotion is a product of an underlying constructive process that is also responsible for overt emotional behavior. The Multilevel Process Theory of Emotions by Leventhal describes adult emotions as complex behavioral reactions that are constructed from a hierarchical multi-component processing system (Leventhal, 1979; Leventhal, 1980; Leventhal and Mosbach, 1983; Leventhal, 1984; Leventhal and Scherer, 1987):

a. Sensory motor level – is a primary emotion generator in response to basic stimulus features,
b. Schematic level – integrates specific situational perceptions with autonomic, subjective, expressive and instrumental responses in a concrete and patterned image-like memory system,
c. Conceptual level – corresponds to social labeling processes.

3.2.1. Sensory Motor Level

The sensory motor or expressive motor level is the basic processor of emotional behavior and experience. It provides the earliest emotional meaning for certain situations. This level consists of multiple components: (a) a set of innate expressive-motor systems and (b) cerebral activating systems. These components are stimulated automatically by a variety of external stimuli and by internal changes of state that do not require deliberate planning. For an example, healthy people will move their hands away when they touch any extremely hot surface without further planning and learning.

Because there is no involvement of the willful planning or learning, the lifetime of the emotional reactions caused at this level may be short and will quickly become the focus for the next level—schematic processing. Action in the facial motor mechanism, as part of the expressive motor system, is the source of the basic or primary emotions of happiness, surprise, fear, sadness, anger, disgust, contempt, and interest (Leventhal, 1979).

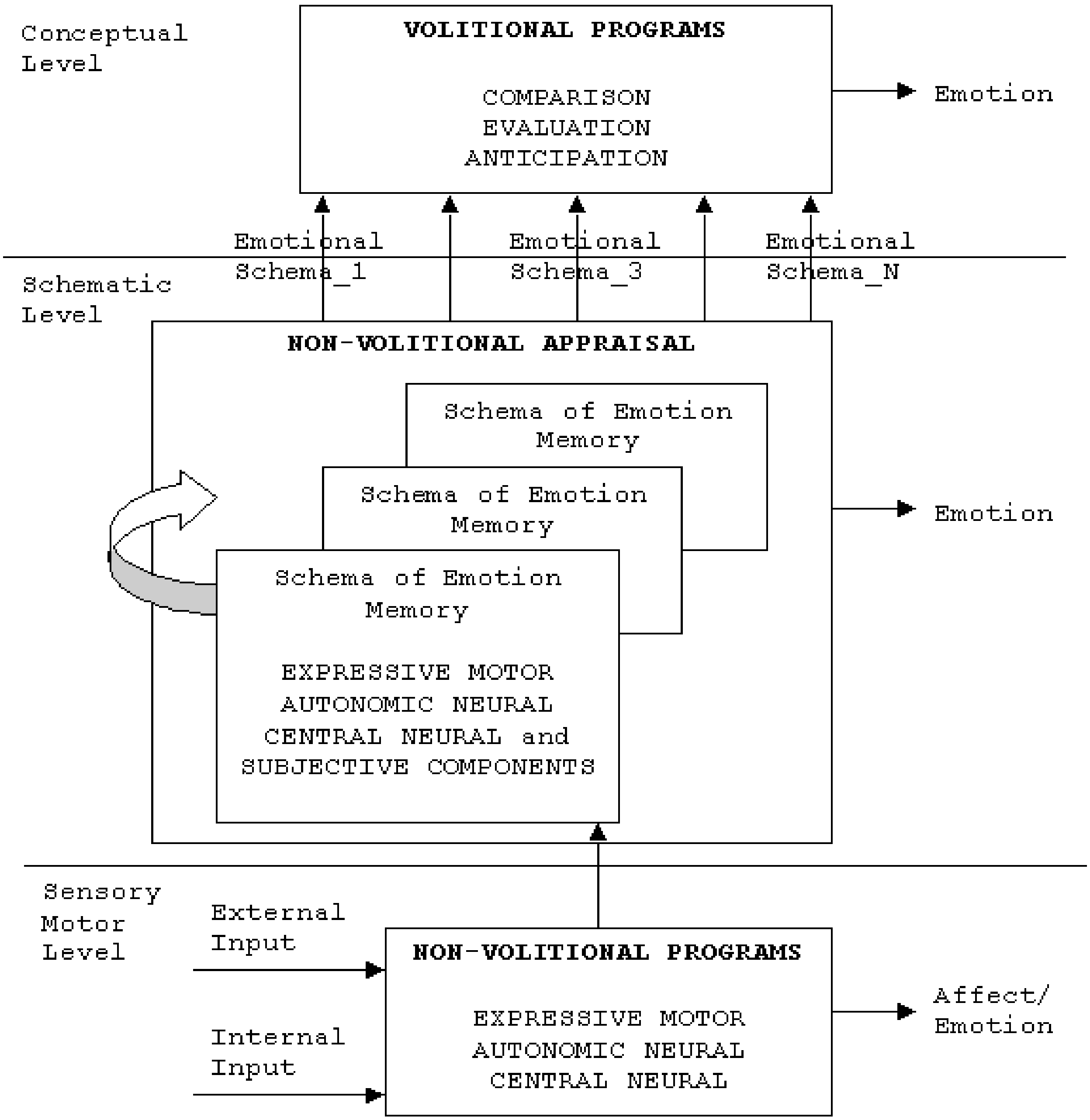


Figure 12: Emotion State Generator (ESG)

3.2.2. Schematic Level

The schematic level integrates sensory-motor processes with prototypes or schemata of emotional situations to create or to structure emotional experiences. But before entering this level, the input needs to be integrated with separate perceptual codes of the visual, auditory, somesthetic (related to the perception of sensory stimuli from the skin), expressive, and autonomic reactions that are reliably associated with emotional experiences. For an example, normal people who like to take a warm shower should not avoid another warm shower in some other new places because a schema of warm water is already in his memory.

Schemata—organized representations of other more elementary codes—are built during an emotional encounter with the environment and will be conceptualized as memories of emotional-experiences. Humans can activate these schemata by activating any one of its component attributes caused by the perception of a stimulus event, by the arousal of expressive behaviors or autonomic nervous system activity, or by the activation of central neural mechanisms that generate subjective feelings. The structure of the schematic memories can be thought as codes, complex categorical units, a network of memory nodes, or perhaps as memory columns that are conceptualized.

Like the sensory motor mechanism, the schematic processing is also automatic and does not require the participation of more abstract processes found at the conceptual level. This schematic level is more complex than the sensory motor level in that it integrates learning processes while building the complexities of schemata. At this level, the lifetime of the emotion behavior is longer than at the previous one.

3.2.3. Conceptual Level

The conceptual level can be thought as the system that can make conscious decisions or choices to some external inputs as well as to internal stimuli, such as stored memories of emotional schemata generated at the schematic level. It is the comparison and abstraction of two or more concrete schemata from the emotional memories with certain concepts that will enable humans to draw conclusions about their feelings to certain events. By comparing and abstracting information from these schemata to the conceptual components—verbal and performance—humans can reason, regulate ongoing sequences of behavior, direct attention and generate specified responses to certain events.

The verbal components do not only represent the feelings themselves but they also communicate the emotional experiences to the subjects (who can also choose to talk about their subjective experiences). On the other hand, the performance components are non-verbal codes that represent sequential perceptual and motor responses. The information contained at this level is more abstract than the schematic memories and therefore the representations can be protected from excessive changes when they are exposed to a new experience and can be led to more stable states. Because this level is volitional, the verbal and performance components can be more sophisticated through active participation of the agent. When performance components are present, for an example, the volitional system can swiftly generate a sequence of voluntary responses to match spontaneous expressive outputs from the schematic level. This volitional system can anticipate emotional behaviors through self-instruction. Both verbal and performance components use a prepositional information network, in which elements are logically related e.g., dog, love, and me.

3.3. Stimulus Evaluation Checks (SECs)

In order to produce emotion for each level above, many researchers have hypothesized that specific emotions are triggered through a series of stimulus evaluations (Scherer, 1984; Scherer, 1986; Weiner, Russell, and Lerman, 1979; Smith and Ellsworth, 1985). In particular, we studied Stimulus Evaluation Checks (SECs) initially proposed by Scherer (1984 and 1986), which were later considered as compatible with Leventhal's Multilevel Process Theory of Emotion (Leventhal and Scherer, 1987).

Indeed, Leventhal and Scherer (1987) proposed stimulus evaluation checks (SECs)—novelty, pleasantness, goal/need conduciveness, coping potential, and norm/self compatibility—for each level of the three levels of the Multilevel Process Theory of Emotion, as shown in Table 2. The integration SECs into the Multilevel Process Theory of Emotion arose because of the dispute between Zajonc, who believed that emotion is primary and independent of cognition, and Lazarus, who believed that emotion is secondary and always dependent upon cognition. In Leventhal's and Scherer's work, they proposed a componential model in which emotions are seen to develop from simpler and reflex-like forms to complex cognitive-emotional patterns that result from the participation of at least two distinct levels of memory and information processing, a schematic and a conceptual level. Continuous stimulus evaluation checks, which evaluate five environment-organism attributes: novelty, pleasantness, goal conductiveness, coping potential, and consistency with social norms and self-related values, activate these systems.

Table 2: Processing Levels for Stimulus Evaluation Checks (adapted from Leventhal and Scherer, 1987).

	Novelty	**Pleasantness**	**Goal/need Conduciveness**	**Coping Potential**	**Norm/self Compatibility**
Conceptual Level	Expectations; cause/effect, probability estimates	Recalled, anticipated, or derived positive-negative evaluations	Conscious goals, plans	Problem solving ability	Self ideal, moral evaluation
Schematic Level	Familiarity; schemata matching	Learned preferences/ aversions	Acquired needs, motives	Body schemata	Self/social schemata
Sensorimotor Level	Sudden, intense stimulation	Innate preferences/ aversions	Basic needs	Available energy	(Empathic adaptation)

Starting from a simplified selection of SECs referred as the *Affective Knowledge Representation* (Lisetti, 2002; Lisetti and Bianchi, 2002), emotions are represented as having many components such as valence, intensity, duration, focality, agency, novelty, controllability, certainty, external norm, internal self standards, facial expression, and action tendency.

As will be explained in the implementation section, because our work is focused on the sensory motor level, which does not need involve deliberative thinking nor learning, we limit the emotion components to the ones we deem relevant at our stage of development already identified by Lisetti (2002):

a. *Valence* [Intrinsic Pleasantness]: *positive/ negative*: is used to describe the pleasant or unpleasant dimension of an affective state. [10]

b. *Intensity/Urgency* [Goal Significance – Urgency]: *very high/ high/ medium/ low/ very low*: varies in terms of degree. The intensity of an affective state is relevant to the importance, relevance and urgency of the message that the state carries.

c. *Focality* [Goal Significance – Focality]: *event/ object*: is used to indicate whether the emotions are about something: an event (the trigger to surprise) or an object (the object of jealousy).

d. *Agency* [Coping Potential – Agent]: *self/ other*: is used to indicate who was responsible for the emotion, the agent itself *self*, or someone else *other.*

e. *Modifiability* [Coping Potential – Control]: *high/ medium/ low/ none*: is used to refer to the judgment that a course of events is capable of changing.

---

[10] The explanations a. through e. use the following format: the name of the component is in italics, followed by Scherer's check equivalency in the square brackets '[ ]'. If the check's equivalency is followed by a dash '-', it is then followed by the sub-system described in (Scherer, 1988). The possible components values are followed after a colon ':' and afterward, the description of the component literally taken from (Lisetti and Bianchi, 2002) follows.

f. *Action tendency*: identifies the most appropriate (suite of) actions to be taken from that emotional state. For example, happy is associated with generalized readiness, frustration with change current strategy, and discouraged with give up or release expectations.

g. *Causal chain*: identifies the causation of a stimulus event associated with the emotion. For example, surprise has these causal chains: (1) something happened now, (2) I did not think before now that this will happen, (3) If I thought about it, I would have said that this will not happen, and (4) Because of this, I feel something. On the other hand, happy has these causal chains: (1) something good happened to me, (2) I wanted this, (3) I do not want other things, and (4) Because of this, I feel good.

As we will show in the Implementation Section, we have integrated our ESG design with the stimulus evaluation checks (SECs) system that is linked into our proposed three-layered architecture model. In our model, the inputs, accepted from our sensors, are fed to a simple perceptual system where raw information is processed into some understandable and interpretable information. For an example, the perceptual system processes raw sonar readings by removing the invalid readings from the valid ones before deciding the drifting rate that will next be used as the input to our ESG.[11]

---

[11] The process of filtering the valid information from the invalid one is explained in the "Implementation" section.

3.4. Schemata

As described in the schematic level above (Figure 12), schemata are the organized representations of more elementary code built during emotional encounters with the environment and conceptualized as memories of emotional-experiences.

The term schema was initially used by Piaget in 1926 (Piaget, 1926) but the schema theory itself was developed several years later by R.C. Anderson, a psychologist (Anderson, 1977). The concept of schema was first brought to the attention of AI roboticists by Michael Arbib (Arbib, 1987) and later on extensively used by Arkin and Murphy for mobile robotics (Arkin and Murphy, 1990), Iberall and Lyons for manipulation (Iberall and Lyons, 1984), and Draper for vision (Draper et al, 1989).

In humans, information stored in memory can be imagined as connections among (or networks of) intertwined schemata. If a person finds new information that is part of schemata, it can be easier for him or her to grasp the concept and ideas of that information. On the other hand, if the new information is not part of schemata, the person needs to learn and insert the new information to the appropriate schemata. Thus, it may be harder for them to understand new information because of this adjustment and necessary learning.

For an example, we can imagine a normal person's schema of a fine dining restaurant, which may be different from a fast food restaurant. In the person's memories, a fine dining restaurant should have connections among buildings (slots and roles: unique decorations, unique lighting, piano, live music, huge aquariums, pillars, chandeliers, etc.), food (slots and roles: eat, appetizers, caviar, foie gras, soup, steak, etc.), drinks (slots and roles: wine, vodka, whiskey, etc.), services (slots and roles: waiters with tuxedos, hostesses with gowns, reservations, etc.), and utensils (slots and roles: complete set of silverware, expensive china, etc.). So when the person enters the fine dining restaurant, she expects to find some, or even all, of the sub-schemata above starting from the greetings by the hostess wearing a black gown to being served by waiters in tuxedos. If the waiter offers them hot dogs, they may be surprised because this kind of food does not (typically) exist in their fine dining restaurant schema. With the absence of the hot dog information, the person will need to learn, adjust her schema of fine dining restaurant, and include hot dogs in the food schema.

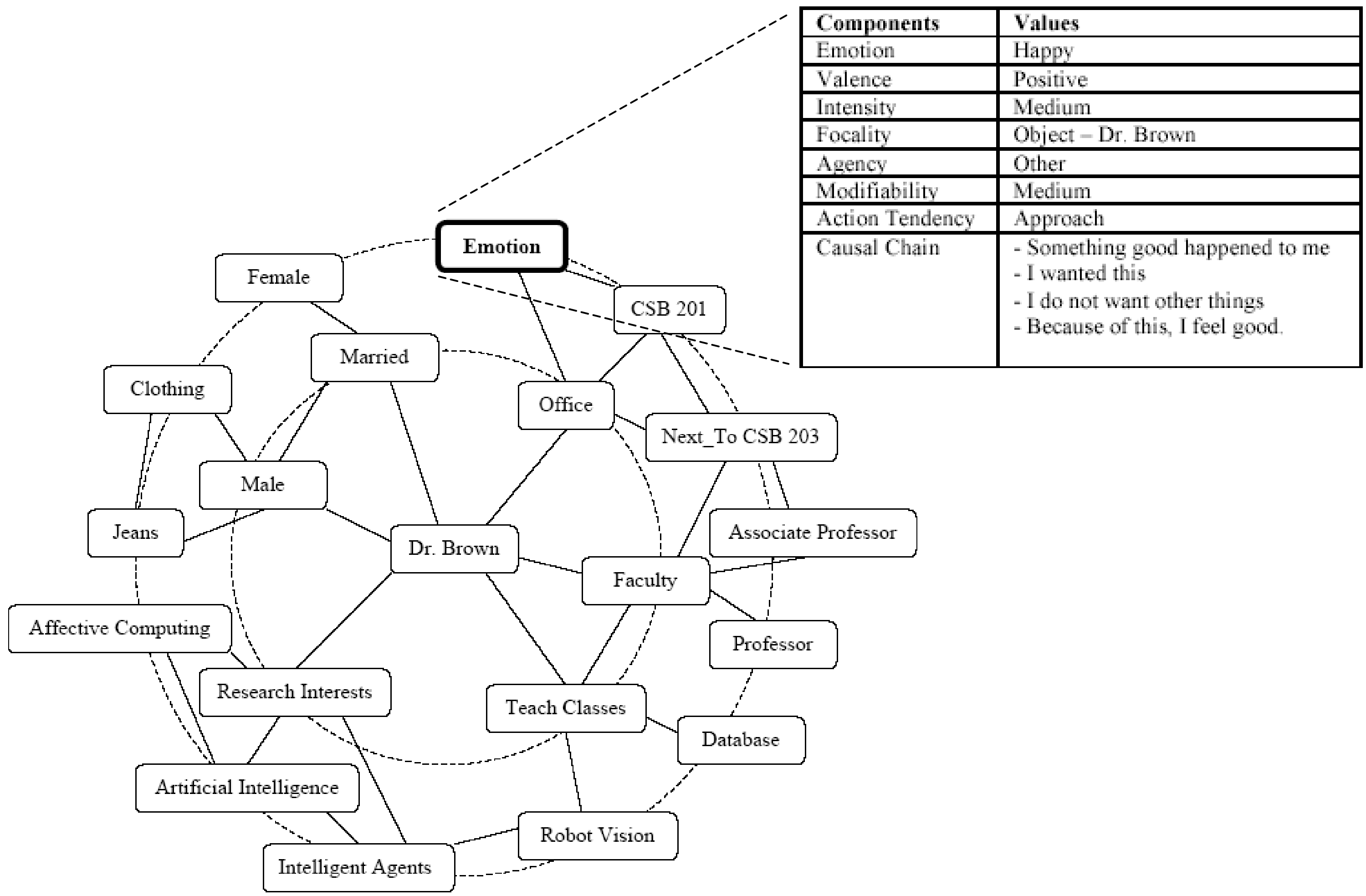
Figure 13[12] shows an example of Petra's complete schemata of Dr. Jones, the Computer Science professor whose office is in CSB 204. The entire information correlated to Dr. Jones is interconnected, including the emotion schema, which is built as a result of SECs from the sensory motor level and explained later in this thesis.

Dr. Jones, a polite and nice professor, always has nice interactions with Petra because he is always available when Petra visits him. He also cooperates when she asks him to face the camera. The interactions with him, as well as his social status—professor—always make Petra happy (Emotion: Happy, Valence: Positive, Intensity: Medium, and Modifiability: Medium). The good interactions also make her want to approach him every day (Action Tendency: Approach). Since this schema is about Dr. Brown, then both focality and agency should point to others (Focality: Other–Dr. Brown and Agency: Other). This emotion schema can also be changed during the interactions with him as well as with the environment. If for some reason Petra always finds many obstacles while navigating to Dr. Brown's office sequentially and repetitively and he does not want to cooperate anymore, her action tendency, as well as other emotion components that always wants to approach him, can be changed internally to start wanting to avoid him.

To continue our effort in emotion modeling, we implement the first layer of the three-layered architecture of ESG in the PeopleBot™. Our efforts are explained further in the next section—Implementation.

---

[12] Each circle represents a schema of an agent and each schema may contain different data structures correlating to appropriate information.



Components	Values
Emotion	Happy
Valence	Positive
Intensity	Medium
Focality	Object – Dr. Brown
Agency	Other
Modifiability	Medium
Action Tendency	Approach
Causal Chain	- Something good happened to me - I wanted this - I do not want other things - Because of this, I feel good.

Figure 13: Example of Petra's Schemata of Dr. Brown

CHAPTER FOUR: IMPLEMENTATION

This chapter focuses on the detailed implementation of our first and second phases where the first one has progressed from having a small robot—Cherry—with no formal emotion mechanism to having a PeopleBot™—called Petra—a much more versatile robot with the Emotion State Generator (ESG), the more formal approach. The topics covered in this chapter include:

- First phase development: Cherry, an office assistant – Section 4.1 contains details on Cherry, an office assistant and an Amigobot™ for our first phase development.
- First phase development: Lola, the robot entertainer – Section 4.2 contains details on Lola, an Amigobot ™ and Cherry's laboratory companion as a robot entertainer.
- Background – Section 4.3 contains details about prior work and reasoning behind the switch from the AmigoBot™ to PeopleBot™.
- Second phase development: Petra, the PeopleBot – Section 4.4 contains details about the robotic hardware/ equipment. Besides explaining the main robot, PeopleBot™, this section also explains the hardware additions, e.g., touch screen, cameras, USB ports, etc. as well as the user-friendly interface.
- Second phase development: Architecture – Section 4.5 contains details on the proposed architecture that links ESG to the perceptual system as well as BSG.
- Second phase development: Sensory Motor Level of the ESG – Section 4.6 contains details on the fuzzy logic implementation of the Sensory Motor Level that includes the Stimulus Evaluation Checks (SECs).

- Second phase development: Emotions – Section 4.7 discusses the emotions that Petra can express at the Sensory Motor Level along with their corresponding facial expressions.
- Second phase development: Multi-Threaded Interface – Section 4.8 contains details on the multithreading implementation that enables Petra to act and behave naturally in real time.

4.1. First Phase Development: Cherry, an Office Assistant

Cherry is a prototype. She is designed to be an office assistant and a tour guide that can be operated in an office environment, specifically the second floor of the Computer Science Building (CSB) at the University of Central Florida.

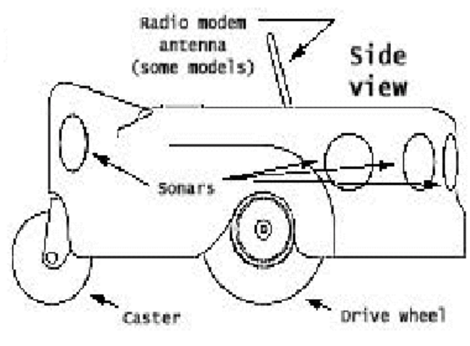
As an office assistant, she has to become a gopher for the faculty and staff. For example, she can be asked to carry Dr. Brown's documents to the copy room, request someone to make copies, and deliver it back to him. She can also deliver a can of Coke to every professor in the department. As a tour guide, she has to introduce the University of Central Florida in general, and particularly the Computer Science Department, to visitors. In introducing the department, she has to describe the research interests of selected professors in front of their office doors.

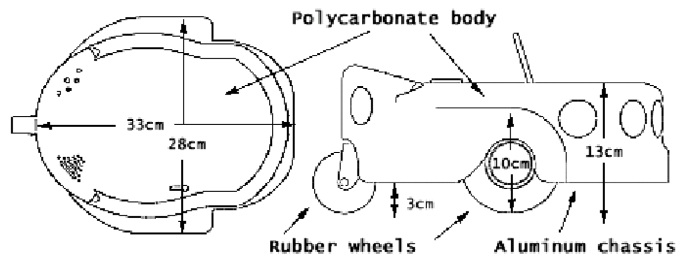
After describing the selected professor, she has to recognize whether the door is open or closed. If it is open, she comes in and requests the professor to stand and face the camera mounted on her. If she can recognize the professor's face, she greets him: "Good morning, Dr. Green." After achieving what she needs to do—greet the faculty—she gets out from the room and changes her strategy. But on the other hand, if she cannot recognize any faces as the corresponding resident of that office, her emotion state changes, explained in the Emotion State Transitions below, after she recursively fails her tasks in performing face recognition for several times. When she reaches the highest level of emotion, anger, she gets out from that office and changes her strategy. On the other hand, if she detects a closed door, her emotion state changes the same way as if she cannot recognize any faces. For both cases, open or closed door, and when her strategy needs to be changed, she requests the users to click on the next faculty or staff to be visited from the point-and-click map. Otherwise, if there is no other request from users, someone can send her back to her home position–CSB 104 (Dr. Lisetti's office suite).

4.1.1. Hardware

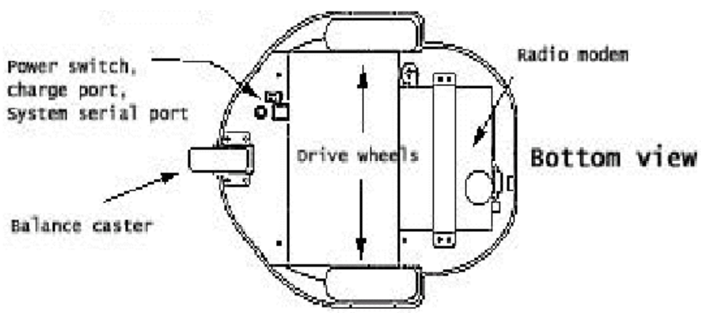
We integrate our works into an AmigoBot™, an intelligent mobile robot, manufactured by the ActivMedia (Activmedia, 2002). AmigoBot™ is capable of autonomous or user defined movement and behavior. It not only has an operating system, but is also packaged with several other programs that allow users to manipulate the robot. The AmigoBot™ is intended for use in areas such as schools, labs, offices, and any place that is wheelchair accessible. The robot is highly maneuverable with 2 rubber tires, each driven by a reversible DC motor, and a rear caster for support. It has a red polycarbonate body that resists damage from running into obstacles. It also has eight sonar sensors that can detect different angles such as 12, -12, 44, -44, 90, -90, 144,

-144 degrees and its size is relatively small (33 cm x 28 cm) with a weight of 3.6 Kg. The side, top and bottom views of this robot are shown in Figure 14.





(a) (b)



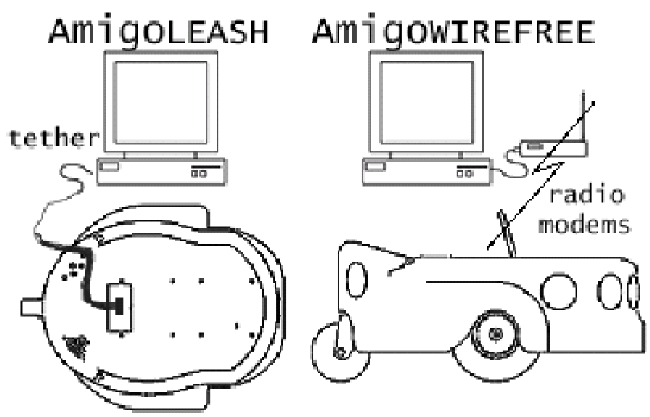
(c)

Figure 14: (a) AmigoBot™ side view,
(b) top view, and (c) bottom view
(adapted from ActivMedia, 2002)

The AmigoBot product also offers two choices of connection types:

(1) direct connection to the robot by using a tether from a PC serial port (Figure 14a). Besides offering a much faster and more reliable method of transferring information, the length of the cable significantly reduces the overall uses that the robot can perform.

(2) wireless capabilities provided by a pair of wireless modems (Figure 14b). One modem is connected to the robot and stored underneath between the wheels and the other is connected to the serial port of a PC. These modems have a range of approximately 300 ft, but considerably less in areas with many walls or sharp turns.

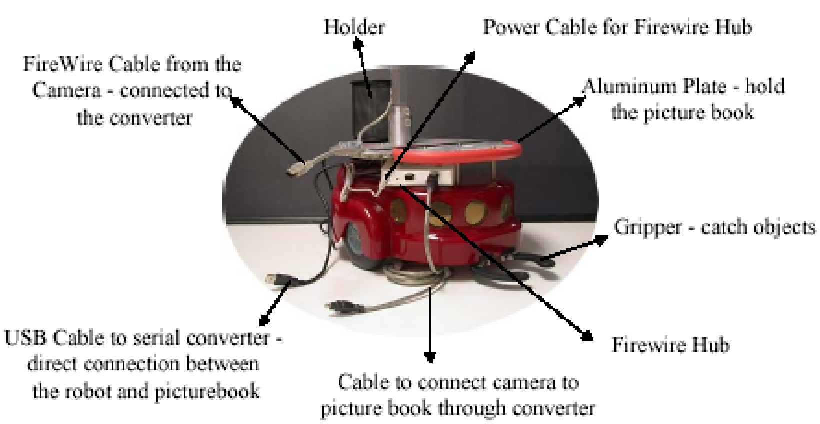


(a) (b)

Figure 15: (a) Direct Connection, (b) Wireless Connection
(adapted from ActivMedia, 2002)

To overcome lack of intelligence in a basic robot (Figure 16a) while looking at its payload limitation (up to 2 lbs), we install a small SONY VAIO PictureBook™ (PCG-CIMV/M)(Sony, 2002), which not only will provide a direct connection to the robot but also additional mobility since it too runs off of battery power. Besides having a PictureBook™ that can boost its capabilities, we also add necessary cables to connect the PictureBook™ to an Orange Micro I-Bot Camera (Orange Micro, 2002), which is mounted at eye-level, to enhance its capabilities to recognize someone (Figure 16b). Finally, due to its limited intelligence, we have to dress it up with some red feathers so users will not have high expectations from it. And to personify it as part of our approach to make the human-robot interaction to be more like the human-human interaction, we name it Cherry (Figure 16c).

(a)



(b)

(c)

Figure 16: (a) The original manufactured AmigoBot™, (b) Cherry's additional hardware components, (c) Cherry with her dress ready to interact with the human beings

4.1.2. Software

ActivMedia provides the ActivMedia Robotics Basic Suite, which contains several programs. They allow users to control their robots using custom-made maps of a building or several rooms, navigate using point-and-click blueprints, or drive using a keyboard or joystick while the robots avoid obstacles.

4.1.2.1. WorldLink

WorldLink, Figure 17, is the basic navigation module that lets users control a robot's actions through point-and-click properties or menus. Through the simulation, the robot's sonar readings are displayed in blue. In addition, its speed and status readings are shown in real time. From this software, users can choose to let the robot wander, drive using point-and-click on a map, or drive with the arrow keys in a keyboard or joystick. The most important feature of the WorldLink is its ability to use the PC as the robot's server. Everyone from around the world can log onto the setup PC, and with permission, they can control the robot and view its movement, audio, and visual input (if the robot is equipped).

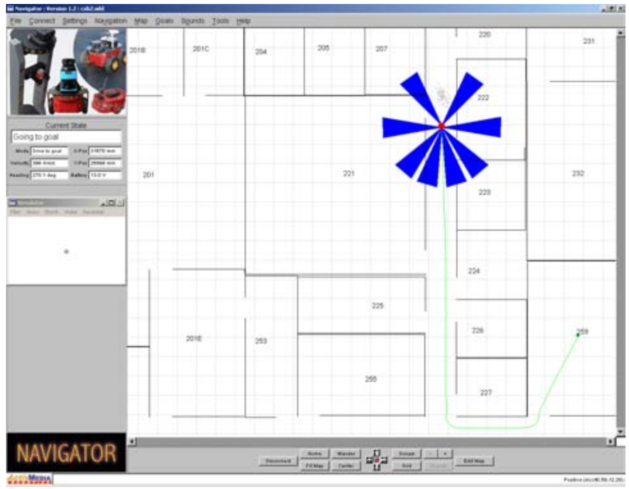


Figure 17: WorldLink (adapted from Activmedia, 2002)

<u>4.1.2.2. Navigator</u>

This software is actually nearly identical to WorldLink. The only difference is that it lacks the web control capability. It still has the same point-and-click features, as well as keyboard or joystick driving capabilities. It, too, utilizes a map for navigation and displays sonar, speed, and status readings.

<u>4.1.2.3. Mapper</u>

This software allows users to create a map of a building, room, or whatever they envision the robot navigating in. Mapper, shown in Figure 18, is essential because other programs, such as WorldLink and Navigator, use a map to direct the robot created in this software. Only walls and large permanent obstacles have to be drawn in, with sketching tools, as the robot is able to use its sonar to navigate around smaller or moving items.

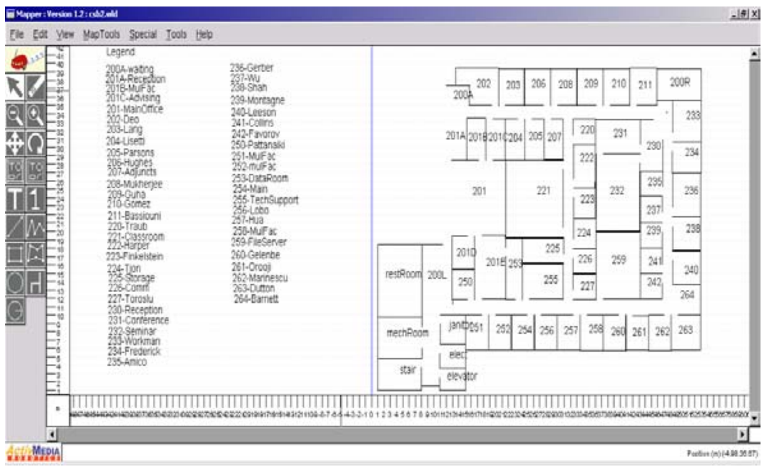


Figure 18: A floor plan model in Mapper

<u>4.1.2.4. Trainer</u>

Using Colbert, a subset of the ANSI C language with a few extensions for robot control, users can write code and create programs for their robot, from simple tasks to more complicated behaviors. The program created can make the robot execute a single command or a series of commands without having to compile. This program is especially useful for novice programmers and for building demos or testing movement portions of larger code.

4.1.2.5. ARIA

ActivMedia Robotics Interface for Application (ARIA) is an object-oriented, robot control applications-programming interface for the ActivMedia Robotics' mobile robots. ARIA, which follows Hybrid Deliberative/Reactive architecture, is a C++- and Java-based language that provides access to the robot server as well as to the robot's sensors and accessories. The classes that comprise ARIA are available and can be used in other C++ or Java codes, provided that it is compiled under Microsoft Visual C++ and Microsoft Visual J++.

4.1.3. Interface

In order to enhance interaction, a Graphical User Interface (GUI) is developed for the robot, which is executed in the PictureBook™. The GUI, shown in Figure 19, integrates several components such as an avatar, an anthropomorphic face, a point-and-click map, a navigational system, a face recognition system, and several other properties, i.e., speech text box, search property, and a live video-capture frame.

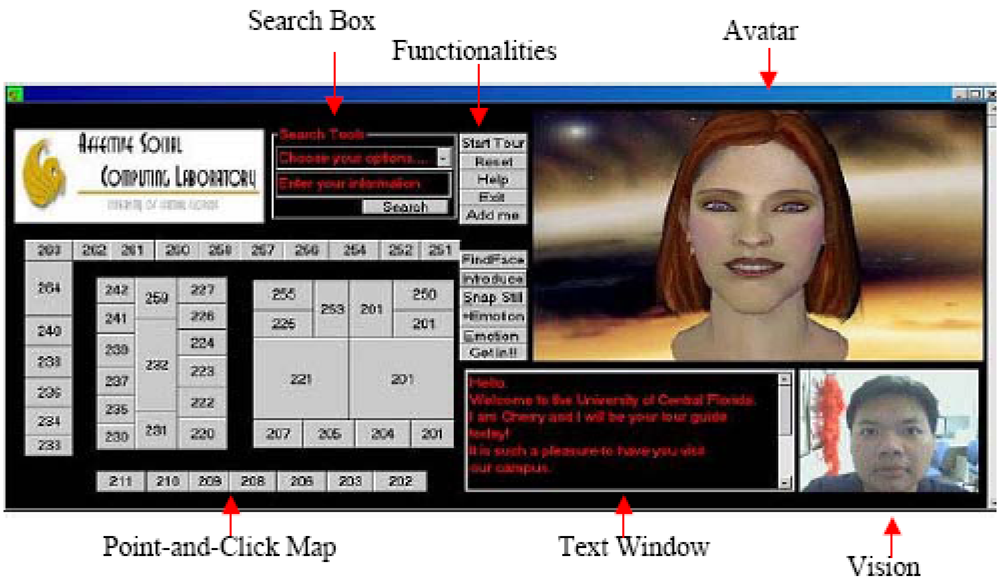


Figure 19: Cherry's Integrated Interface

4.1.3.1. Avatar

It has been shown by Lisetti and Schiano (2000) that since Darwin (1898), many researchers' interests have been correlated with the movements of the face primarily, which are associated with the expressions of inner emotional states. These observations motivated the integration of facial expressions elements as a way to show the inner state of the agent. With this consistent anthropomorphic face, it can be expected that humans would understand the robot's emotional state better and much faster.

With advances in graphics, the anthropomorphic faces have become more natural and human-like. Cherry's avatar is created with the Haptek's People Putty (Haptek, 2002). Cherry is designed to be a twenty something young woman who is both attractive and capable of displaying and influencing others with some facial expressions. Besides having natural facial expressions, skin tone, and etc., the anthropomorphic face is able to mimic human movements e.g., random head and eye movements, and lip movements, as well as the human voice[13].

4.1.3.2. Point-and-Click Map

In the interface, the map of the office suites is designed as a point-and-click environment (Figure 20). The buttons are assigned to a specific room in certain x- and y- coordinates and each number corresponds to a suite number belonging to the faculty and staff members. For each room on the map, it stores some information about the faculty or staff who works in that room, i.e., name, title, and research interests, which can be used to greet residents of the room and/or introduce them to the users. Clicking a certain button from the map triggers her movement from her current position to the chosen one by executing the navigation system that allows her to move forward between the aisles, represented as the spaces between the x- and y- coordinates of the walls.

[13] Haptek Player allows the users to choose several different options of voices (male, female, robots, and voice simulations in space, in a stadium, on a telephone, and whispering). To be more natural and appropriate, a female voice was chosen as the default.

4.1.3.3. Navigational System

Through the point-and-click map in Figure 20, users can select several faculty and staff to visit. These button-clicking events trigger the execution of the navigational system algorithm. All rooms and hallways are represented with the x- and y-coordinates, measured in millimeters[14]. Having all these points, the robot can move through the aisle between the x- and y-coordinates of the walls and intersections.

Figure 20: Map of Faculty and Staff Suites

Figure 21 shows an example in which Cherry (X) needs to move from a point A (80, 100) to a point B (120, 200). From the starting point, she needs to move between (55, 195) and (75, 195) until she reaches any intersection with the y-coordinate less than 200. In this case, Cherry continues her navigation through the T-junction (75, 185). Since the x-coordinate of point A is smaller than point B, she needs to make a right turn in the junction. Afterward, she can continue her journey in the aisle between (75, 195) and (85, 195), until she reaches point B and makes a left turn.

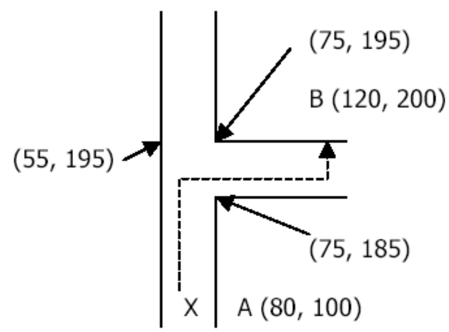


Figure 21: Case Study of the Navigation System

4.1.3.4. Face Recognition System

For Cherry to have intelligence to recognize someone, a face recognition system is added by integrating the FaceIt™ technology designed by Identix (Identix, 2002) to the interface. This system starts its execution when Cherry needs to perform face recognition. Through the Micro Orange I-Bot camera, a bitmap image, called temp.bmp, is captured, which is temporarily stored in the directory DIR. Since the technology can only process a .jpeg file, it then needs to be

[14] Millimeter is used to ease the distance and movement calculations that parallel with the usage in ARIA, the C-based language used in the robots

converted using the JPEGDLL.dll file located in the C:\WINNT\system32 directory. After the conversion is done, it then becomes an input to the FaceIt™ technology.

Along with our algorithm that traverses a certain directory[15] DIR and picks a .jpeg file, this technology compares the biological similarities, such as cheekbone and eye characteristics, between these two files that then gives a matching degree. If this point is greater than or equal to eight, which shows high similarities, the comparison algorithm stops, and then checks the filename corresponding to the image. Otherwise, the searching is continued until the end of DIR. If the algorithm finds a matching person, it then needs to find the information related to the image. If it does not find anyone, it then stops and Cherry tells users that she does not recognize that person (NOT_RECOGNIZE), which then follows the transitions mentioned in the Emotion State Transition section below.

To find the information of the corresponding person, the algorithm needs to change the image filename's extension to the text filename's extension, only for the sake of the name. For example, to find Andreas Marpaung's information, the filename AndreasMarpaung.jpeg, which has the highest matching degree, is changed into AndreasMarpaung.txt. With the text filename, the algorithm can find and open it, from the same directory DIR, and then read the content of the file. The first line of the file has the name of that person (NAME) while the second one has the social status information (STATUS) (faculty, staff, student ε STATUS). So based on this social

---

[15] A database of images and text files consists of the names, social statuses and social interactions (are used for fuzzy logic) of correlated images. Each pair of image (*.jpeg) and text file (*.txt) is named after the person's name, but with different extension. For example, the image file is called AndreasMarpaung.jpeg, while the text file is called AndreasMarpaung.txt

status, Cherry then can choose an appropriate greeting (GREETING) at certain time (TIME) (morning, afternoon, evening ε TIME), according to the following rules:

If STATUS == "faculty" then GREETING = "Good (TIME), Dr. (NAME)"

Elseif STATUS == "staff" then GREETING = "Good (TIME), (NAME)"

Elseif STATUS == "student" then GREETING = "Hey dude, what's up!"

The flowchart of the face recognition system is shown in Figure 22.

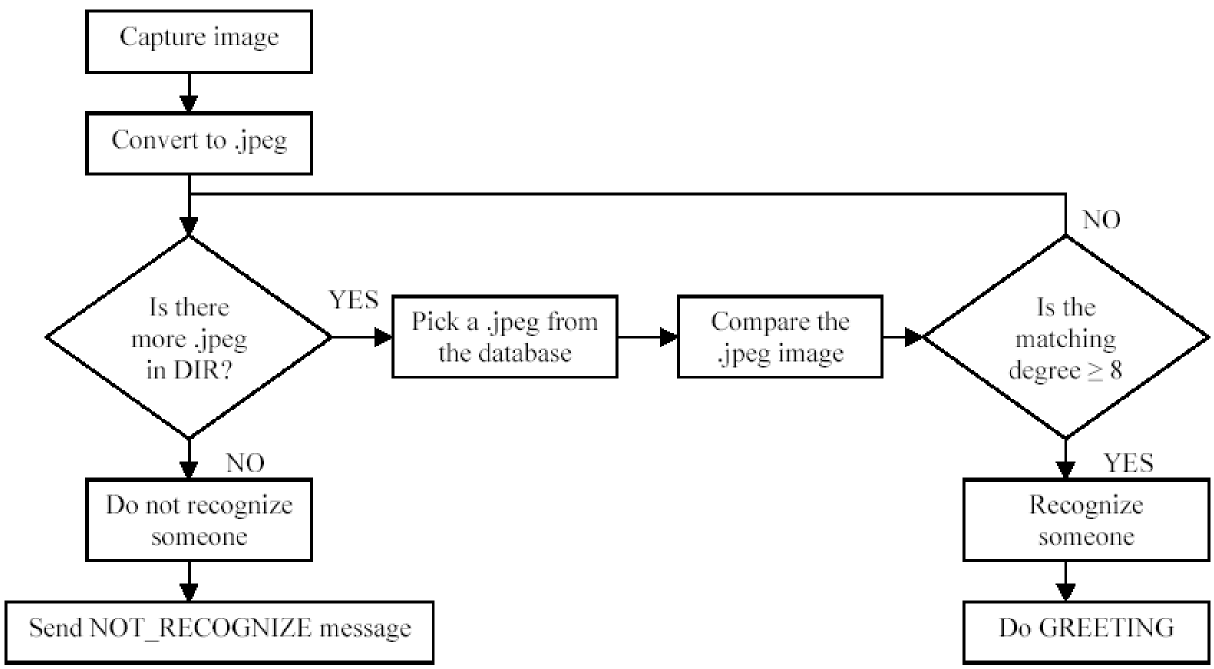


Figure 22: Face Recognition Flowchart

4.1.3.5. Other Properties

Realizing that the interactions will not always happen in a quiet environment, users may find it hard to comprehend what Cherry has said in a noisy situation. In order to eliminate this misunderstanding and to enhance the interaction in such circumstances, a *speech text box* was added to the interface. Through this box, users are able to read and understand what she actually says to them in any environment.

Another property that the GUI has is the *search box*. This box allows users to perform a simple searching algorithm to find the room numbers and residents based on known information.

A live video-capture frame is also integrated into the GUI. Through "Cherry's eyes," users can see everything that Cherry sees in her world. With the existence of this frame, the faculty and staff who are asked to stand and face the camera can also ensure that they are inside the camera view or not. If they are not, they can align themselves so that the face recognition algorithm can receive the right image input.

4.1.4. Emotion State Transitions

As mentioned above, Cherry's emotions move from one state to another because of the occurrences of external events: recognize an open (or closed) door and recognize (or not recognize) people. With these external inputs, her external belief changes from happy to neutral, neutral to frustrated, and frustrated to angry (shown in Figure 23).

If Cherry feels frustrated for a certain period of time, her internal belief state changes from confident to discouraged. When she starts with the happy external belief and confident internal belief, she thinks that she is capable of achieving her tasks, but after repetitive failures, her internal belief changes from confident to discouraged, in which she feels that she cannot achieve any other tasks and it remains the same until she receives some positive stimuli.

In order to show her emotion, each state is associated with a certain facial expression shown in Figure 24.

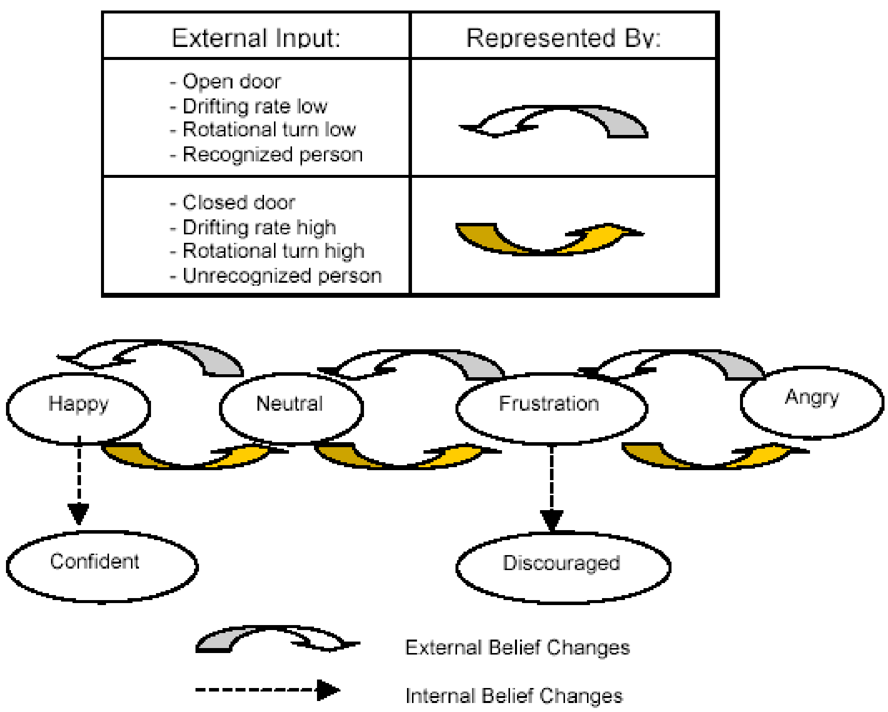
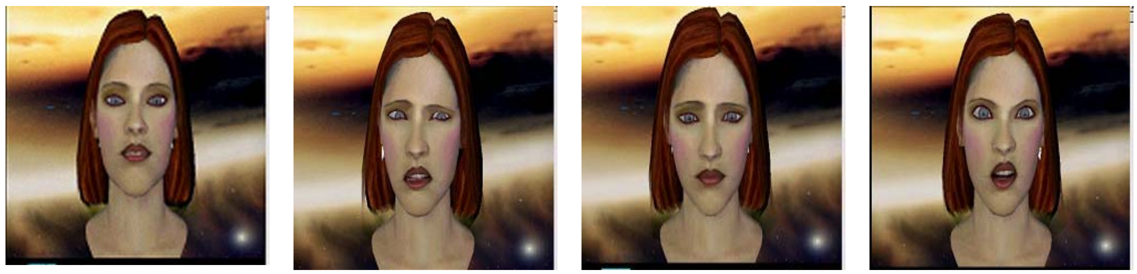


Figure 23: Cherry Finite State Emotions Transition



(a) (b) (c) (d)

Figure 24: Cherry's facial expression
(a) Happy, (b) Frustrated, (c) Discouraged, and (d) Angry

4.2. First Phase Development: Lola, the Robot Entertainer

In contrast to Cherry as an office assistant and a tour guide, Lola, her laboratory companion, is the same AmigoBot™ but with a different interface and purpose. She is designed to have some capabilities to dance for and entertain audiences by featuring some state-of-the-art multimedia developments integrated to a robotic platform (Marpaung, Brown, Lisetti, 2002).

As an entertainer, Lola's main capability is to dance in front of the audiences. She has seven selected songs to dance along to the predefined routines in her audio collection: (1) "All You Need is Love," (2) "Mr. Roboto," (3) "Supermodel," (4) "Lola," (5) "La Vie en Rose," (6) "Little Red Corvette," and (7) "We are Alive."

4.2.1. Hardware

The hardware that is used for Lola is almost the same as the one in Cherry. The only difference is that Lola has a pair of speakers powered by two AA batteries so the battery lifetime of the robot is not affected. Since Lola needs to move around, the weight of the speakers, which is connected to the speaker output-jack of the PictureBook™, is also considered. These light speakers enable users to raise the volume, compared to a small one on the PictureBook™, so that they can hear to what she says as well as to the dance songs that are played.

Besides the speakers, to distinguish Lola from Cherry, we dressed Lola with a red and a black feathers as well as a black beret. The overall appearance of Lola is shown in Figure 25.

Figure 25: Lola's overall appearance

4.2.2. Interface

Cherry's interface is the backbone to Lola's, as shown in Figure 26. Most of Cherry's elements are eliminated except the avatar, the speech textbox, the live video-capture, and the ASCL logo.

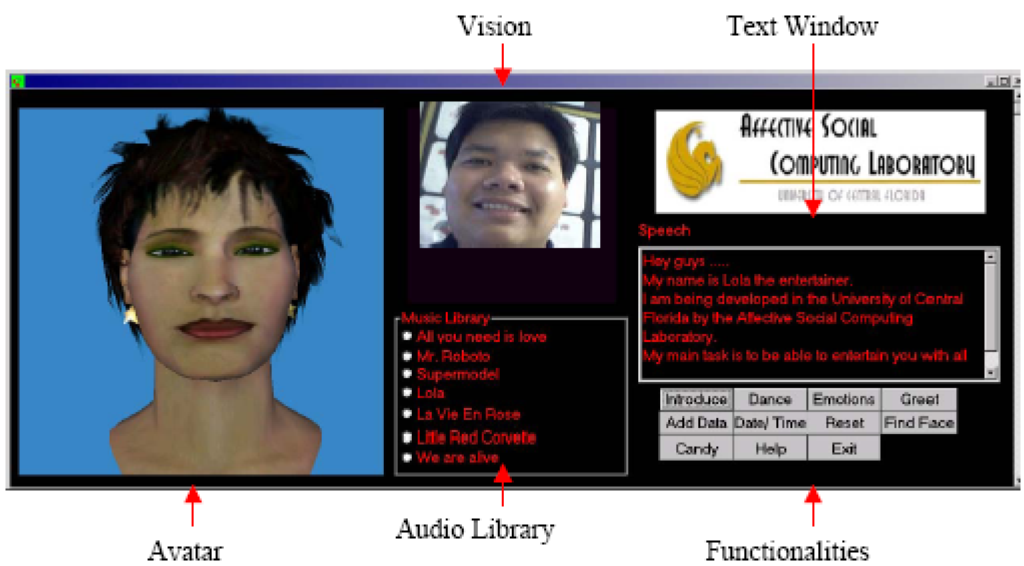


Figure 26: Lola's integrated interface

From the interface, users are able to select a song for Lola by clicking on a radio button. When someone chooses a song, Lola soon executes her predefined dance routine by moving forward, backward, turning around, and flowing to the rhythm.

Besides being capable of dancing and showing her internal state, Lola is also capable of:

a. introducing herself. By clicking the *Introduce* button she can explain everything about her personality. For example, "Hi! I am Lola. I am an AmigoBot™

programmed in the Affective Social Computing Laboratory at the University of Central Florida. I like to sing, dance, and do so many things."

b. explaining the procedures to operate her system properly as well as the interface features that she has to offer.
c. greeting someone she recognizes after executing the face recognition algorithm. Otherwise, she has to apologize to them for not having their data in her database.
d. telling the date and time. With this feature, she is also able to greet someone properly, e.g., "Good morning" at 8 a.m. instead of "Good night."
e. resetting her current emotion-state. Someone needs to click on the *Reset* button when Lola reaches the anger state so her emotion-state can be downgraded back to her happy state.
f. exiting the interface when users click the *Exit* button.

4.2.3. Emotion State Transitions

Lola is also capable of displaying her current emotion states that follows the same emotion transitions as in Cherry's but are triggered differently. Unlike Cherry, whose emotions are triggered by the open or closed door and the recognized or unrecognized person, Lola's emotions are triggered because of the repetitive tasks that need to be performed following these rules:

If PERFORMITR == 0 then EMOTION-LIKE-STATE = "Happy"

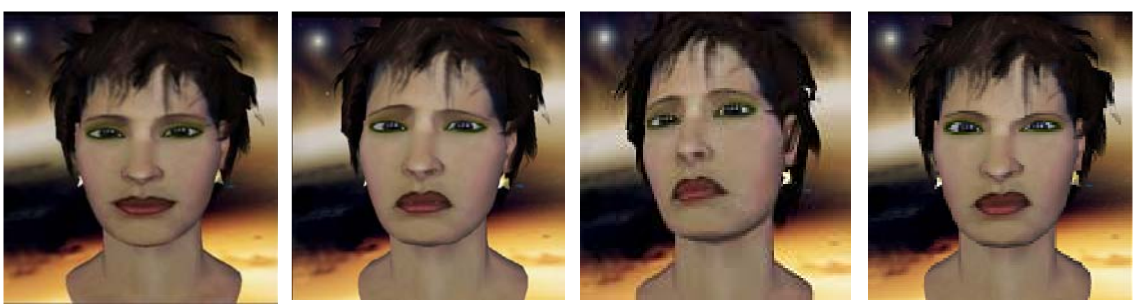
Elseif PERFORMITR == 2 then EMOTION-LIKE-STATE = "Frustrated"

Elseif PERFORMITR == 4 then EMOTION-LIKE-STATE = "Discouraged"

Elseif PERFORMITR == 6 then EMOTION-LIKE-STATE = "Angry"

where PERFORMITR is the frequency of the same task performed sequentially represented by numerical values between 0 and 6, and EMOTION-LIKE-STATE is her emotion state (Happy, Frustrated, Discouraged, Angry ε EMOTION-LIKE-STATE). For example, she feels frustrated if she needs to introduce herself more than twice. After introducing herself more than four times, she feels discouraged because she thinks that users never listen to her explanation. And finally, when she introduces herself for the sixth time, she reaches her highest emotion state—anger. When she reaches this state, she refuses to perform any commands from users until someone clicks the "Reset" button that sets her emotion back to the default--happy.

Just like Cherry, she is also capable of showing her emotion-state through her facial expression, as shown in Figure 27.



(a) (b) (c) (d)

Figure 27: Lola's facial expression
(a) Happy, (b) Frustrated, (c) Discouraged, and (d) Angry

4. 3. Background

The emotion-modeling project starts when we think of having an emotional social robot to enhance Human-Robot Interaction (HRI). Our initial idea is to have a small AmigoBot™, Cherry, to become a tour guide for the visitors and an office assistant for the faculty and staff members on the second floor of the Computer Science Department. After having additional hardware installed and dressing her up, a survey is conducted for our social informatics approach, to rate her acceptability as a tour guide and gopher (Lisetti, Brown, Alvarez, and Marpaung, 2004). One of the results shows that the ideas of having a robot to express emotions are warmly accepted by the participants. With these results from the survey, the prototype code in Cherry is transferred to a bigger and much more versatile PeopleBot™ that we name Petra.

In the original design, a formal mechanism for producing emotion is not technically implemented—there is only a finite state machine of emotion transitions that is only linked to face recognition. As it has become clear from the survey that a social robot with emotions and expressions could prove useful. Our second phase approach is documented in this thesis. In our second phase, the first level of the Multilevel Process of Emotion—sensory motor level—is implemented using fuzzy logic. The stimulus captured by the sensors—sonar, camera for face recognition, camera for navigation and obstacle avoidance system—is fed to the Emotion State Generator that will modify the intensity of emotion represented as scripts or schemata.

4.4. Second Phase Development: Petra, the PeopleBot

4.4.1. Hardware

In order to transfer the initial code, from Cherry to Petra, some new hardware is added to the PeopleBot™. The additions are: a Smart Display™ touch screen, a camera for face recognition (at eye-level), a camera for navigation and obstacle avoidance (at floor-level), a wireless antenna for the touch screen, and some USB ports. This sub section explains the additions in more detail.

4.4.1.1. Robot

With the previous weight and height limitations, AmigoBot™ is exchanged for the PeopleBot™. PeopleBot™ is designed to have better interactions with humans. In contrast to Cherry having her interface at the floor level, the new PeopleBot™, with its human height, enables the interface to be at the appropriate height so users can point and click the interface comfortably. PeopleBot™ comes with twenty-four sonars within its two rings (sixteen sonars on the bottom ring and eight sonars on the top one) that can improve the navigational and obstacle-avoidance algorithm compared to Cherry, an AmigoBot™ with eight sonars. Besides sonars, PeopleBot™ also has a gripper, which can be used as her limited hand, with four degrees of freedom–move-up, move-down, opened, and closed.

PeopleBot™ also has an on-board full-sized PC with a Hitachi H85 processor that currently uses Windows XP. Originally, the robot comes with both Windows 2000 and RedHat Linux as its possible operating systems. Since the transferred code is compatible with Windows 2000, its operating system is originally chosen for Petra. After adding the touch screen that can only be operated in the XP environment, however, it becomes apparent that the operating system needs to be switched from Windows 2000 to Windows XP. Through this on-board PC, the interface can be executed and displayed on the screen. In the original design, the PeopleBot™ does not come with any USB ports. To accommodate the camera and wireless antenna for the touch screen, USB cables are connected to the internal PCI port.

4.4.1.2. LCD Touch Screen

With the initial motivation in mind—to enhance HRI—our hardware is designed to be as user-friendly as possible. To support user-friendliness, the display media is changed from the small VAIO PictureBook™ in Cherry, into a bigger Smart Display™—DesXCape 150DM (Microsoft, 2003). This 15-inch Smart Display™ is the latest LCD touch screen technology that connects to and accesses a nearby Windows XP Professional-based desktop remotely. With this technology, the keyboard and mouse can be eliminated and the interface can be operated with a stylus. This touch screen, placed on the top of the robot, is connected to the robot's internal PC (a nearby Windows XP Professional-based desktop), using 802.11 wireless antenna plugged into a USB port.

With the light touch screen that uses its own battery power, the interface can be executed through the internal PC while moving around and performing the robot's tasks safely.

4.4.1.3. Cameras

Due to the fixed position of the pan-tilt camera that comes with the PeopleBot, we have to add another two cameras that can be mounted in the appropriate positions. Since the PeopleBot does not come with a firewire board, the Micro Orange I-Bot that uses the firewire connection is switched to the USB camera. For Petra's purposes, a LogiTech camera mounted at floor-level is used for the navigational and obstacle avoidance system. Another USB camera, Intel (Intel, 2002), used to capture the image for the face recognition system, is mounted at eye-level.

4.4.2. Graphical User Interface

In order to enhance interaction, a Graphical User Interface (GUI) is developed for the robot and displayed through the user-friendly touch screen. The GUI as well as the integrated hardware are shown in Figure 28. Through a survey (Lisetti, Brown, Alvarez, and Marpaung, 2004), we find that users are pleased with the usability and functionalities of the interface. We also find that they are fond of Cherry's avatar. With these main reasons along with its more complex tasks than Lola, which can support the design of ESG, we decide to keep using Cherry's interface and functionalities in the PeopleBot™. The interface integrates several components such as the avatar, an anthropomorphic face, a point-and-click map, emotion changing progress bars, several algorithms (navigational, vision and obstacle avoidance, and face recognition systems), several help menus, e.g., speech text box, search properties, and start-at-room option, and two live-capture frames.

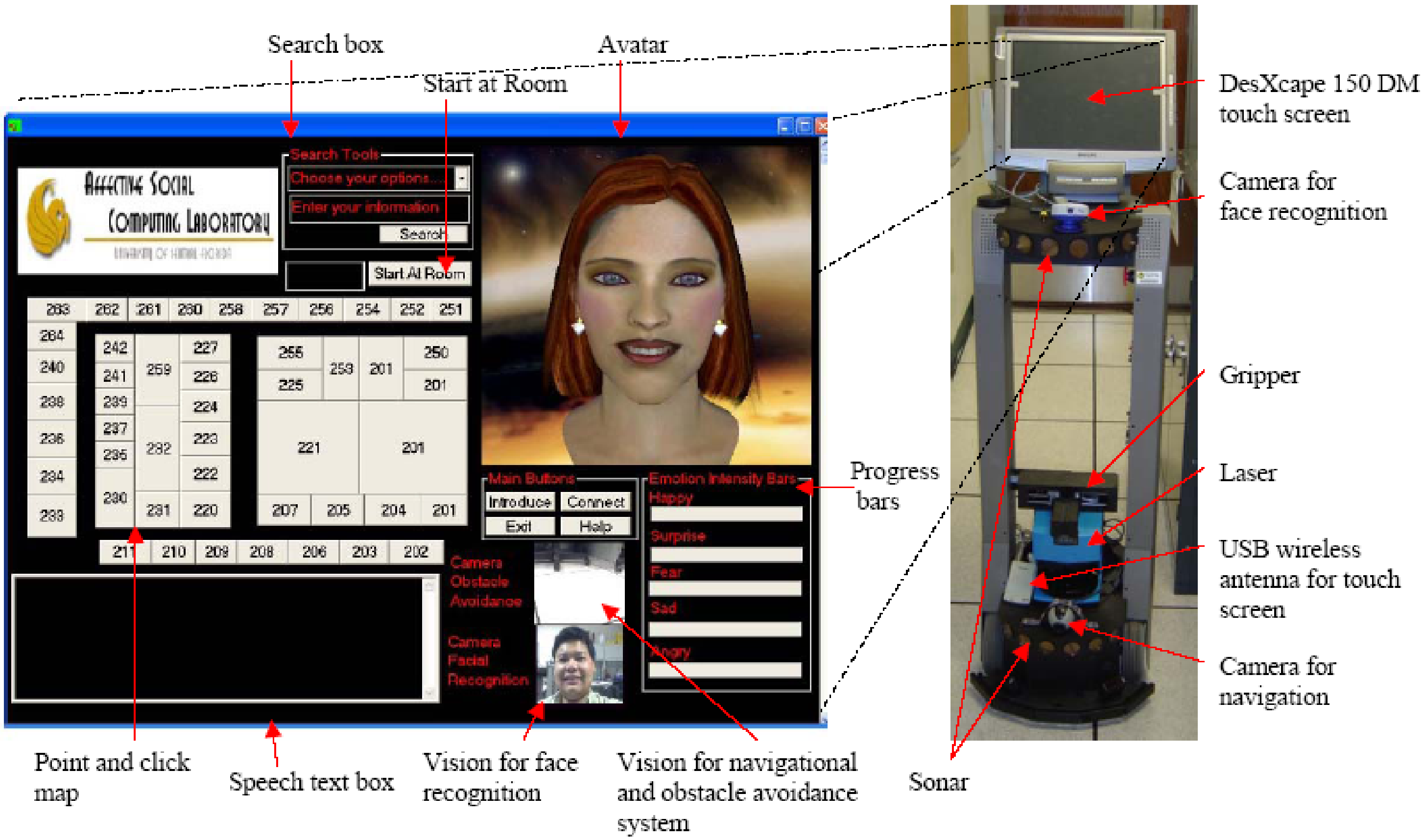


Figure 28: Petra's complete interface and hardware

4.4.2.1. Avatar

In the survey (Lisetti, Brown, Alvarez, and Marpaung, 2004), the participants were asked whether they prefer Cherry's avatar to Lola's and we find out that the majority of them prefer to use Cherry's. With these majority voices, we still keep using Cherry's face as Petra's.

4.4.2.2. Progress Bars

In addition to displaying the emotion-like state through the avatar, progress bars are added. Through these bars, users are able to know the real-time changes and certain emotions affected at certain times by certain stimuli. For each cycle, the emotion intensity calculated by the Emotion State Generator changes the values of the progress bars based on the OR-mapping (described further in the "From Fuzzy Logic to Emotion-Like States" section below).

4.4.2.3. Speech Text Box, Search Box, and Start-At-Room Button

Both the speech text box and search box have the same functionalities as Cherry's. The only functionality added is the *start-at-room button*. Currently, to ease the navigational system, the robot is stationed at a certain room (room 204—Dr. Lisetti's office) and sets that room as her default starting point. With this functionality, users have the flexibility to reset the starting point for current execution. Besides resetting the starting point, with this button, users are also able to set the current position of the robot should it be needed. For example, if for some reasons the navigation and obstacle avoidance system ends up with some discrepancies in distance, i.e., travels too far, users can move her manually to a closer room, instead of bringing her back to her home position—CSB 204, and continue the navigation from that room after resetting the starting point.

4. 5. Second Phase Development: Architecture

Figure 29 below shows the architecture that is developed in Petra, which consists of (1) Perceptual System, (2) Emotion State Generator (ESG), and (3) Behavior State Generator (BSG). All systems, except the ESG (which can be found in the Multilevel Process Theory of Emotion in the Approach section), are explained in more detail. Its implementation is explained in the sensory motor level section below.

4.5.1. Perceptual System

The perceptual system is designed to be a simple system; it can filter and convert raw information abstracted through the sensors into some meaningful information that can be used as an input to the ESG and BSG. Currently, as can be seen from Figure 29, Petra has three different sensors or stimuli—sonars, a camera for face recognition, and a camera for navigation and obstacle avoidance. We describe how each sensor is programmed and used.

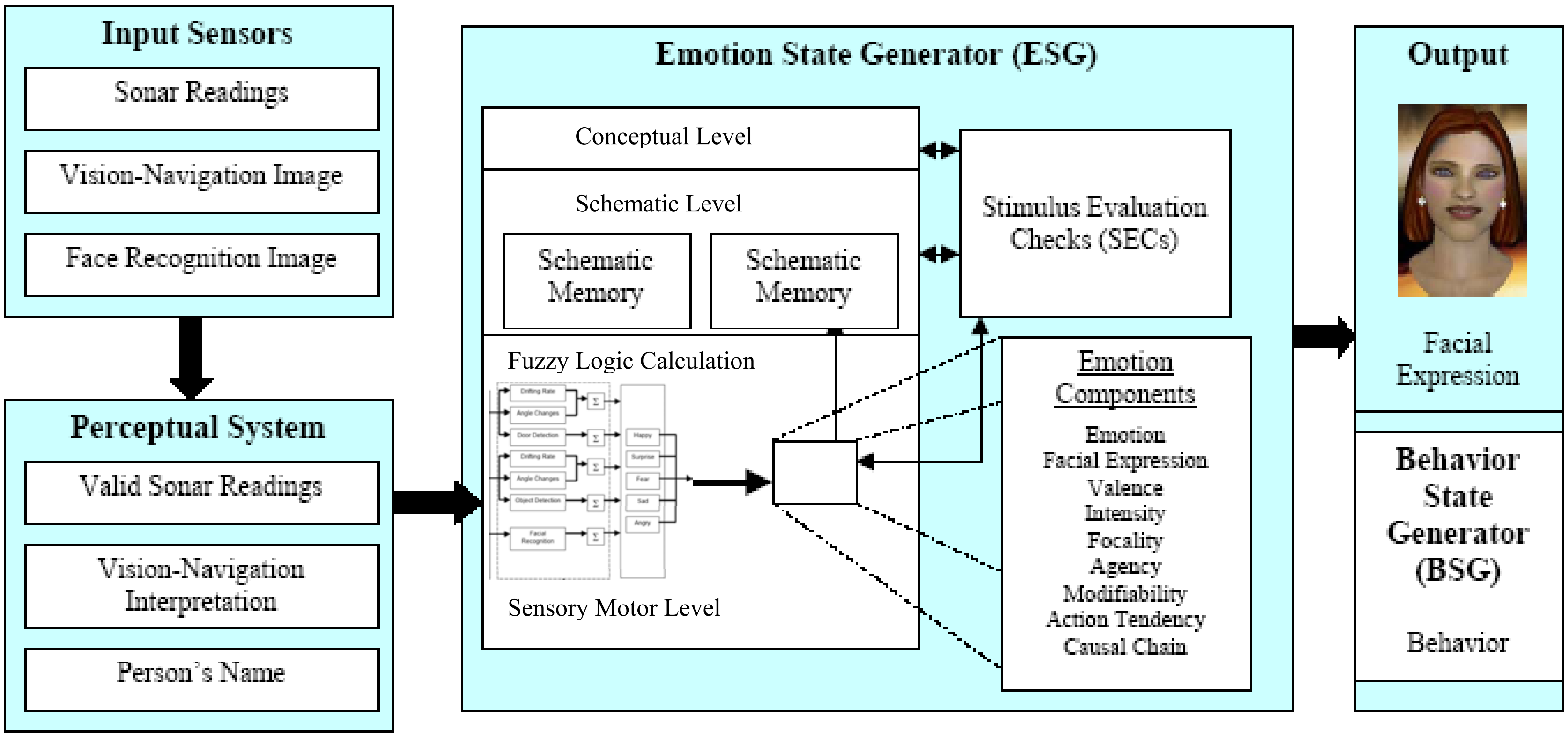


Figure 29: Three-layered Architecture

4.5.1.1. Sonars

In our robot, the sonars can sense any objects by processing the readings received after firing up the sonar. Since the hallway information is known and determined, i.e., the width of the hallway aisle, the fixed distance between the door and robot, etc, the readings can tell whether there is any objects or obstacle nearby or not.

In order to make our implementation uniform, a cycle is set to 1000 mm navigational distance for the perceptual system to accept any inputs. For each cycle, the sonar performs readings within 200 mm increments to keep the accuracy of the abstracted information. These readings, stored in an array, are obtained from two different sonars at the bottom ring—the leftmost and the rightmost. So for each cycle, the perceptual system receives five different raw readings, which can be either valid or invalid[16].

Out of these five readings, the invalid ones must be thrown away so the system only stores valid information to be processed further. The reading is invalid if the sum of the left and the right sonar readings is extremely more or extremely less than the distance between the aisles (1,500 mm for our case). And vice versa, the reading is valid if the sum is around 1,500 mm. After cleaning up the invalid data, the perceptual system restores this valid information into an array and sends it to the ESG and BSG to be processed further.

---

[16] The validity and invalidity of the readings can be decided based on the robot's current position. The total right-most and left-most sonar' reading, when the robot is in the middle of the hallway, equals to around 1500 mm is valid but invalid if it equals to 5000 mm.

4.5.1.2. Camera for face recognition

Due to Petra's current limitation on the on-board PC and the processing power needed for face recognition to be performed, the perceptual system receives an input from this camera only when the robot stops, asks the person to be recognized to come closer, and captures the facial image. When the image is captured, it is sent to the face recognition algorithm that consists of the image selection from her library and FaceIt technology by Identix. To find any matching, the algorithm compares the similarities, represented as the matching degree, between the captured and selected images from the library. When the algorithm finds the highest matching degree, Petra greets that person by name and social status otherwise she ignores that person[17]. The result of recognized face or not is also sent to the ESG to be processed further[18].

4.5.1.3. Camera for obstacle avoidance

For each cycle, the camera captures an image and sends it to the vision algorithm, as part of the perceptual system. In this system, the images are smoothened and edged by a canny edge detector that eliminates some unnecessary lines due to the shadows. Since the main goal of this algorithm is to have some lines with some degrees of diagonality in order to detect the wall edges and/or obstacles, the algorithm eliminates the vertical edges from the image and leaving it

[17] At current stage, Petra does not have any intelligence to remember new person's name that she encounters.

[18] Currently, the information of the recognized person is not sent to/implemented in the ESG. But in the future, this information may be integrated with the sensory motor level and needed for the schematic and/ or the conceptual level where further learning and information processing can be done.

with the non-vertical edges. With these retained edges, if there is no obstacle, the system can detect the vanishing point by calculating the farthest point in the hall. Otherwise, if there are obstacles, the system can detect the closer lines, with certain lengths, that are certain distances away from the vanishing point. The results of these findings—the coordinates of the vanishing point or the starting and ending points of the obstacles—are sent to both ESG and BSG for further processing.

4.5.2. Behavior State Generator (BSG)

According to Murphy (2000), a behavior is "a mapping of sensory inputs to a pattern of motor actions, which then are used to achieve a task." This behavior can be further divided into three broad categories: *reflexive behaviors,* "hardwired" responses to the stimuli so the response time can be shortened; *reactive behaviors,* learned behaviors that then can be produced without conscious thought; and *conscious thought,* deliberate behaviors. Out of these three types, the reflexive behavior was chosen for this project because the sensory motor level responds to the stimuli.

After the perceptual system filters the stimuli, the system sends them to the ESG. They are then forwarded to the BSG. Currently, the BSG is built as a simple mechanism that helps smoothen the navigation by centering the robot in the middle of the aisle and avoiding any simple obstacles. Through these outputs from the perceptual systems, the robot can execute different behaviors depending on the input source. Each behavior state is described below:

INIT: Reset the emotion and the progress bars to the default setting—happy—and the initial position to room 204.

STAY_CENTER: Center herself between the aisles to avoid the walls.

AVOID_LEFT_WALL: Move to the right to avoid the left wall. This behavior is triggered should course correction calculated by sonar and/or vision be needed.

AVOID_RIGHT_WALL: Move to the left to avoid the right wall. This behavior is triggered should the course correction be needed.

WAIT: Wait for a fixed period of time when the face recognition algorithm cannot recognize anyone or the door is closed.

4.6. Second Phase Development: Sensory Motor Level

Inspired by FLAME (El Nasr, 2002), the sensory motor level is implemented with fuzzy logic, described in this section, in order to resolve uncertainty about the perceptual system. In short, fuzzy logic is "a logical system that generalizes classical two-valued logic for reasoning under uncertainty" (Yen and Langari, 1999).

Out of many different techniques and methods to measure fuzziness, the emotion-like states[19] are modeled with the Takagi Sugeno Kang (TSK) model (Yen and Langari, 1999). H. Takagi and M. Sugeno first introduce the TSK model as an additive rule model and later on, continued by M Sugeno and K.T Kang on the identification of this type of fuzzy models. We

[19] From this section forward, emotion-like states are interchangeably used with the emotion states due to our effort to "model" emotion, not to "create" emotion.

choose this model because it is simple and easy to understand and can also reduce the number of rules required.

In general, rules in the TSK model have the form of:

IF x_1 is A_{i_1} and … x_r is A_{i_r}

THEN $y = f_i (x_1, x_2, \ldots, x_r)$

$= b_{i\_0} + b_{i\_1} * x_1 + \ldots + b_{i\_r} * x_r$

where x_r is the condition's variable, A_{i_r} is the condition itself, f_i is the linear model, and b_{i_j} ($j = 0, 1, \ldots, r$) are real-valued parameters. From these rules, in order to derive the final conclusion, the model aggregates the conclusion of multiple rules using an inference analogous to the weighted sum that is shown below

$$y = \frac{\Sigma\, \alpha_i * f_i (x_1, x_2, \ldots, x_r)}{\Sigma\, \alpha_i}$$

where α_i is computed as:

$$\alpha_i = \min (\mu_{Ai\_1}(a_1), \mu_{Ai\_2}(a_2), \ldots, \mu_{Ai\_r}(a_r))$$

where μ_{Ai} is the matching degree of input $x_1 = a_1, x_2 = a_2, \ldots, x_r = a_r$.

In short, based on fuzzy logic, the sensory motor level is modeled as shown in Figure 30. This level receives valid outputs from the perceptual system and processes them further in order to produce emotion-like states. This level accepts: (1) from the sonar, it receives the course correction degree (drifting rate and angle changes) and door detection; (2) from the camera for navigational and obstacle avoidance system, it receives the course correction degree (drifting rate and angle changes) and object detection, and (3) from the camera for face recognition, it receives the person's name as the result whether the chosen person is recognized or unrecognized.

We now explain each of the three subsystems in details.

4.6.1. Fuzzy Logic from Sonar

From the perceptual system, the valid sonar information is processed further in order to give the x-y drifting rate (Δ drift) and the angle change (Δ angle) by comparing the previous information with the current ones. Both the drifting rate and the angle change have five fuzzy sets, which in the program are represented by sequential numerical values between 1 (the lowest) and 5 (the highest) as follows:

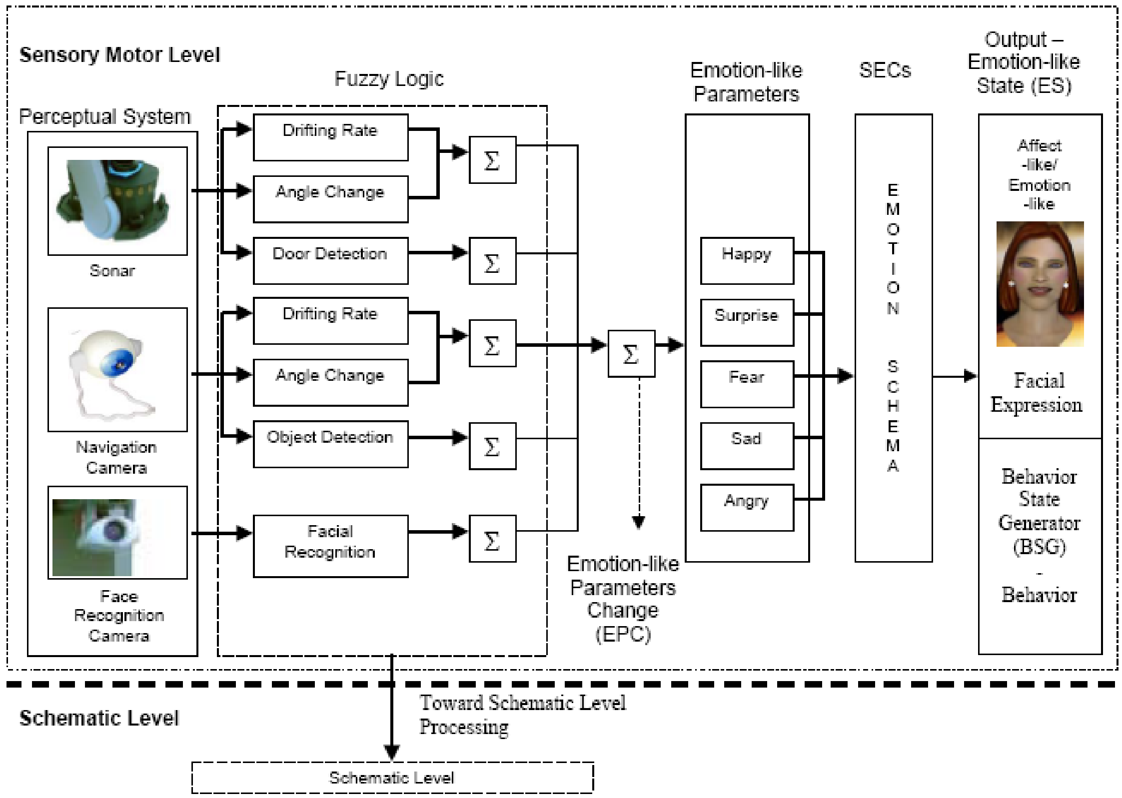


Figure 30: Sensory motor level Architecture

a. Drifting rate fuzzy sets (F_{drift_sonar})

a.1. LowestDriftRate—drifting between 0 mm and 100 mm
a.2. LowerDriftRate—drifting between 100 mm and 200 mm
a.3. MediumDriftRate—drifting between 200 mm and 300 mm
a.4. HigherDriftRate—drifting between 300 mm and 400 mm
a.5. HighestDriftRate—drifting higher than 400 mm

b. Angle changes fuzzy sets (F_{angle_sonar})

b.1. LowestAngleChange—Δ angle is between 0^O and 18^O
b.2. LowerAngleChange—Δ angle is between 18^O and 36^O
b.3. MediumAngleChange—Δ angle is between 36^O and 54^O
b.4. HigherAngleChange—Δ angle is between 54^O and 72^O
b.5. HigherAngleChange—Δ angle is between 72^O and 90^O

Following TSK model, the emotion's intensity caused by drifting rate and angle change can be calculated as follows:

1. Compute x_{drift_sonar} and y_{angle_sonar}:

IF Δ drift is $x_{drift\_sonar\_input}$ THEN $x_{drift\_sonar} = drift(x_{drift\_sonar\_input})$

IF Δ angle is $y_{angle\_sonar\_input}$ THEN $y_{angle\_sonar} = angle(y_{angle\_sonar\_input})$

where $x_{drift\_sonar\_input}$ is the drifting rate in millimeters , $y_{angle\_sonar\_input}$ is the angle change in degrees, x_{drift_sonar} is the drifting rate fuzzy value, y_{angle_sonar} is the angle change fuzzy value, $drift(x_{drift\_sonar\_input})$ ε[20] (F_{drift_sonar}), and $angle(y_{angle\_sonar\_input})$ ε (F_{angle_sonar})

2. Compute $\alpha_{drift\_and\_angle\_sonar}$:

According to TSK model, the $\alpha_{drift\_and\_angle\_sonar}$ is calculated as follows:

$$\alpha_{drift\_and\_angle\_sonar} = \min\left(\mu_{drift\_sonar}(x_{drift\_sonar\_input}),\ \mu_{angle\_sonar}(y_{angle\_sonar\_input})\right)$$

But because the drifting rate and angle change come from one source—from sonar—the μ_{drift_sonar} is the same as μ_{angle_sonar}.

$$\begin{aligned}\alpha_{drift\_and\_angle\_sonar} &= \mu_{drift\_sonar}(x_{drift\_sonar\_input}) \\ &= \mu_{angle\_sonar}(y_{angle\_sonar\_input}) \\ &= \frac{\text{frequency of valid readings in one cycle}}{\text{frequency of readings in one cycle}} * 100\%\end{aligned}$$

[20] ε means the element of.

where there are five readings in one cycle that contain both valid and invalid readings and

$0 < \mu_{drift\_and\_angle\_sonar} \leq 1$

3. The emotion-like state's intensity caused by drifting rate and angle change is

$$y_{drift\_and\_angle\_sonar} = \frac{\alpha_{drift\_and\_angle\_sonar} * (x_{drift\_sonar} + y_{angle\_sonar})}{\alpha_{drift\_and\_angle\_sonar}}$$

:

For door detection, we have two possible cases—open and closed—with a corresponding fuzzy value for each case:

c. Door detection ($B_{door\_detection}$)

c.1. OpenDoor—detected as an open door
c.2. ClosedDoor—detected as a closed door

At this moment, Petra does not have any capabilities to detect whether the door is widely or half open. Thus, with this limitation, we can only assume that the door is widely open, even though it is half open. Thus, it always sets the matching degree to 1. In other words,

$$\alpha_{door\_detection} = \mu_{door\_detection}(x_{door}) = 1$$

where $x_{door} \in B_{door\_detection}$, which is either open, represented by 5, or closed, represented by –5, in the program. And, finally, the output for the door detection (y_{door}) is:

$$y_{door} = \frac{\alpha_{door} * x_{door}}{\alpha_{door}}$$

Thus, the emotion-like state's intensity caused by the sonar is:

$$y_{sonar} = \frac{[\alpha_{drift\_and\_angle\_sonar} * (x_{drift\_sonar} + y_{angle\_sonar})] + [\alpha_{door} * x_{door}]}{\alpha_{drift\_and\_angle\_sonar} + \alpha_{door}} \quad \text{[Eq. 1]}$$

4.6.2. Fuzzy Logic from Vision

The fuzzy logic mechanism from vision is the same as the one from sonar since both of them work together for the navigational and obstacle avoidance system. Both of them can calculate the drifting rates and angle changes but the sonar recognizes doors while the vision detects any obstacle in front of the robot. Both drifting rate and angle change have five fuzzy sets represented as sequential numerical values between 1, the lowest, and 5, the highest, as follows:

a. Drifting rate fuzzy sets (F_{drift_vision})

a.1. LowestDriftRate—drifting between 0 mm and 10 mm
a.2. LowerDriftRate—drifting between 10 mm and 20 mm
a.3. MediumDriftRate—drifting between 20 mm and 30 mm
a.4. HigherDriftRate—drifting between 30 mm and 40 mm
a.5. HighestDriftRate—drifting higher than 40 mm

b. Angle changes fuzzy sets (F_{angle_vision})

b.1. LowestAngleChange—Δ angle is between 0^O and 18^O
b.2. LowerAngleChange—Δ angle is between 18^O and 36^O
b.3. MediumAngleChange—Δ angle is between 36^O and 54^O
b.4. HigherAngleChange—Δ angle is between 54^O and 72^O
b.5. HigherAngleChange—Δ angle is between 72^O and 90^O

Following the TSK model, the emotion-like state's intensity caused by drifting rate and angle change from the camera for obstacle avoidance can be calculated as follows:

1. Compute x_{drift_vision} and y_{angle_vision}:

IF Δ drift is $x_{drift\_vision\_input}$ THEN $x_{drift\_vision} = drift(x_{drift\_vision\_input})$

IF Δ angle is $y_{angle\_vision\_input}$ THEN $y_{angle\_sonar} = angle(y_{angle\_vision\_input})$

where $x_{drift\_vision\_input}$ is the drifting rate in millimeters , $y_{angle\_vision\_input}$ is the angle change in degrees, x_{drift_vision} is the drifting rate fuzzy value, y_{angle_vision} is the angle change fuzzy value, $drift(x_{drift\_vision\_input})$ is the drifting function that gives drifting fuzzy value and ε (F_{drift_vision}), and $angle(y_{angle\_vision\_input})$ is the angle change function that gives angle change fuzzy value and ε (F_{angle_vision})

2. Compute $\alpha_{drift\_and\_angle\_vision}$:

According to TSK model, the $\alpha_{drift\_and\_angle\_vision}$ is calculated as follows:

$$\alpha_{drift\_and\_angle\_vision} = \min\left(\mu_{drift\_vision}(x_{drift\_vision\_input}),\ \mu_{angle\_vision}(y_{angle\_vision\_input})\right)$$

But because the drifting rate and angle change come from one source—the camera for obstacle avoidance—the μ_{drift_vision} is the same as μ_{angle_vision}.

$$\begin{aligned}\alpha_{drift\_and\_angle\_vision} &= \mu_{drift\_vision}(x_{drift\_vision\_input}) \\ &= \mu_{angle\_vision}(y_{angle\_vision\_input}) \\ &= \frac{\text{frequency of valid readings in one cycle}}{\text{frequency of readings in one cycle}} * 100\%\end{aligned}$$

where there are five images in one cycle and $0 < \mu_{drift\_and\_angle\_vision} \leq 1$

3. The emotion's intensity caused by drifting rate and angle change is

$$y_{drift\_and\_angle\_sonar} = \frac{\alpha_{drift\_and\_angle\_sonar} * (x_{drift\_sonar} + y_{angle\_sonar})}{\alpha_{drift\_and\_angle\_sonar}}$$

For object detection, we have two possible cases—obstacle and no_obstacle—with a corresponding fuzzy value for each case:

c. Object detection ($B_{object\_detection}$)

c.1. Obstacle—detect an obstacle in front of the robot
c.2. No_obstacle—detect no obstacle in front of the robot

The same case as the door detection, Petra does not have any capabilities to recognize the shape or the size of the obstacle. Thus, it always sets the matching degree to 1. In other words,

$$\alpha_{object} = \mu_{object}(x_{object}) = 1$$

where $x_{object} \in B_{object}$, which is either obstacle, represented by 5, or no_obstacle represented by –5, in the program. And finally the output for the object detection (y_{object}) is:

$$y_{object} = \frac{\alpha_{object} * x_{object}}{\alpha_{object}}$$

Thus, the emotion's intensity caused by the camera for obstacle avoidance is:

$$y_{obstacle} = \frac{[\alpha_{drift\_and\_angle\_vision} * (x_{drift\_vision} + y_{angle\_vision})] + [\alpha_{object} * x_{object}]}{\alpha_{drift\_and\_angle\_vision} + \alpha_{vision}} \quad \text{[Eq. 2]}$$

4.6.3. Fuzzy Logic for Face Recognition

Besides accepting inputs from both sonar and camera for obstacle avoidance, the camera mounted to Petra's platform also captures an image to be processed further by the Identix technology. The image is captured when Petra requests the chosen person to stand in front of her.

For face recognition, we have two possible cases—recognized and no_recognized—with a corresponding fuzzy values for each case:

Face recognition detection ($B_{face\_recognition}$)

1. Recognized—recognize the chosen person
2. No_recognized—not recognize the chosen person

Following TSK model, the emotion's intensity caused by the face recognition can be calculated as follows:

1. Compute $x_{face\_recognition}$:

IF face_recognition is $x_{face\_recognition\_input}$ THEN $x_{face\_recognition}$ = face_rec($x_{face\_recognition\_input}$)

where $x_{face\_recognition}$ is the face recognition fuzzy value and face_rec($x_{face\_recognition\_input}$) ε ($B_{face\_recognition}$), which in the program is represented by 5, if recognized, or -5, if not.

2. Compute $\alpha_{face\_recognition}$:

For the camera for face recognition, the $\alpha_{face\_recognition}$ is calculated as:

$$\alpha_{face\_recognition} = \min(\mu_{face\_rec}(x_{matching\_degree}), \mu_{face\_rec}(y_{social\_status}), \mu_{face\_rec}(z_{social\_interaction}))$$

where $\mu_{face\_rec}(x_{matching\_degree})$ is the matching degree function that takes the matching degree value calculated by the FaceIt™ technology, between 0 and 10, and divides it by 10. Thus

$0 < \mu_{face\_rec}(x_{matching\_degree}) \leq 1$. $\mu_{face\_rec}(y_{social\_status})$ is the social status function whose value is between 0 and 1 that shows the social status of the chosen person (stored in the text file explained in section 4.1.). At current implementation, faculty is set to 0.7, staff to 0.5, and student to 0.3. $\mu_{face\_rec}(z_{social\_interaction})$ is the social interaction function whose value is between 0 and 1 (stored in the text file explained in section 4.1.). Currently, it is set to 1. In the future, this value can be changed according to Petra's fondness for interacting with a certain person whose value ranges between 0, which shows that the person does not interact nicely to Petra, and 1, which shows good interactions.

Thus, the emotion-like state's intensity caused by the camera for face recognition is:

$$y_{face\_recognition} = \frac{[\alpha_{face\_recognition} * x_{face\_recognition}]}{\alpha_{face\_recognition}} \quad \text{[Eq. 3]}$$

After processing all inputs with fuzzy logic shown in equations [1] to [3], the emotion-like state's intensity affected or emotion-like parameters change (EPC) is calculated as follows:

$$EPC = y_{sonar} + y_{obstacle} + y_{face\_recognition} \quad \ldots\ldots \text{[Eq. 4]}$$

$$= \frac{\{[\alpha_{drift\_and\_angle\_sonar} * (x_{drift\_sonar} + y_{angle\_sonar})] + [\alpha_{door} * x_{door}]\} + \{[\alpha_{drift\_and\_angle\_vision} * (x_{drift\_vision} + y_{angle\_vision})] + [\alpha_{object} * x_{object}]\} + \{[\alpha_{face\_recognition} * x_{face\_recognition}]\}}{\{\alpha_{drift\_and\_angle\_sonar} + \alpha_{door}\} + \{\alpha_{drift\_and\_angle\_vision} + \alpha_{object}\} + \{\alpha_{face\_recognition}\}}$$

4.6.4. From Fuzzy Logic to Emotion-Like States

The calculated emotion-like parameters change, as shown in equation 4 above, determines the changes in the emotion-like parameters: happy, surprise, fear, sad, and angry. The emotion-like parameters are represented by numerical values that can either be increased or decreased on each cycle based on the OR-mapping[21] shown in Table 3 below based on the following function:

$$EP\ (\text{emotion-like}, t) = EP\ (\text{emotion-like}, t-1) \pm EPC$$

where EP (emotion-like, t) is the function that gives the change for each emotion-like for current cycle t, EP (emotion-like, t-1) is the same function for the previous cycle t-1, emotion-like ε {happy, surprise, fear, sad, angry}, and EPC is the calculated emotion-like parameter change.

After calculating all the emotion-like parameters, the final emotion-like state (ES) is calculated based on the following rule:

$$ES(t) = \max \{EP(\text{happy}, t), EP(\text{surprise}, t), EP(\text{fear}, t), EP(\text{sad}, t), EP(\text{angry}, t)\}$$

Finally, before Petra expresses a certain emotion-like state, each emotional experience needs to go through Stimulus Evaluation Checks (SECs) to build a schema that is stored in the memory for further use. In SECs, the schema with all of emotion components, described in the Approach section and shown in Table 4 below, is assigned to an appropriate value.

---

[21] The OR-mapping is a logical mapping in which one of conditions is (or both conditions are) true in order to make the entire statement true. For example, if A is false and B is true, then A OR B = true.

Table 3: The OR-mapping of the emotion-like parameters

Parameter	Increased if	Decreased if
Happy	- Small to Medium-small value of the processed information from sonar or vision - Open door - Someone recognized	- Medium to Large value of the processed information from sonar or vision - Closed door - Someone not recognized
Surprise[22]	- Large value of the processed information from sonar or vision (on the first detection only)	- The robot is in the happy state
Fear	- Large value of the processed information from sonar or vision (medium repetition)	- The robot is in the happy state
Sad	- Medium to Medium-large value of the processed information from sonar or vision - Closed door - Someone not recognized	- Small to Medium-small value of the processed information from sonar or vision - Open door - Someone recognized
Angry	- Large value of the processed information from sonar or vision (high repetition) - Closed door (repetitively) - Someone not recognized (repetitively)	- Small to Medium-small value of the processed information from sonar or vision - Open door - Someone recognized

Table 4: Emotion Components

No.	Emotion Components	Possible Component Values	Component Descriptions
1.	Valence	Positive/Negative	Used to describe the pleasant or unpleasant dimension of an affective state.
2.	Intensity/Urgency	Very high/ High/ Medium/ Low/ Very low	The intensity of an affective state is relevant to the importance, relevance and urgency of the message that the state carries.
3.	Focality	Event/ Object	Used to indicate whether the emotions are about something: an event (the trigger to surprise) or an object (the object of jealousy)
4.	Agency	Self/ Other	Used to indicate who was responsible for the emotion, the agent itself *self*, or someone else *other*.
5.	Modifiability	High/ Medium/ Low/ None	Used to refer to the judgment that a course of events is capable of changing.
6.	Action Tendency	Varies on emotions	Identifies the most appropriate (suite of) actions to be taken from that emotional state.
7.	Causal Chain	Varies on emotions	Identifies the causation of a stimulus event associated with the emotion.

[22] To show surprise, when the processed information from sonar or vision is large on first detection, this emotion is prioritized.

Besides assigning the components, each schema is also associated with the current object of the emotion such as a walking person or an obstacle. Table 5 shows an example of a surprise schema associated with an unexpected walking person when she navigates from room 201 to room 204.

Table 5: An example of surprise schema

Components	**Values**
Emotion	Surprise
Valence	Negative
Intensity	Very High
Focality	Object
Agency	Other – walking person from 201 to 204
Modifiability	High
Action Tendency	Avoid
Causal Chain	- Something happened now - I did not think before now that this will happen - If I thought about it, I would have said that this will not happen - Because of this, I feel surprised

A sudden appearance of a person in the navigation image is detected as an obstacle that can slow down the navigation process due to the course correction that needs to be performed should the person remain in the navigation image on the next cycle. Thus intensity is very high and the action tendency is to avoid potential obstacles. Since the face cannot be detected at a far distance, the valence is negative. And at current cycle, the modifiability is set to high because she "thinks" that she will soon be able to perform the obstacle avoidance to change the course event. The creation of a surprise schema, which is stored in the memory, can become an input for her decision making in the future.

4.7. Second Phase Development: Emotions

As mentioned in the previous section, in order for users to perceive Petra's internal state better and faster, an anthropomorphic face was added to our interface. At current implementation, Petra can display five different facial expressions, shown in Figure 31 (a-e), associated with the emotions expressed in the sensory motor level and corresponding to her different inner states. Out of these five expressions, Petra only picks one expression based on the calculated ES(t). For example, if ES(t) = EP(happy, t), a happy facial expression is chosen.

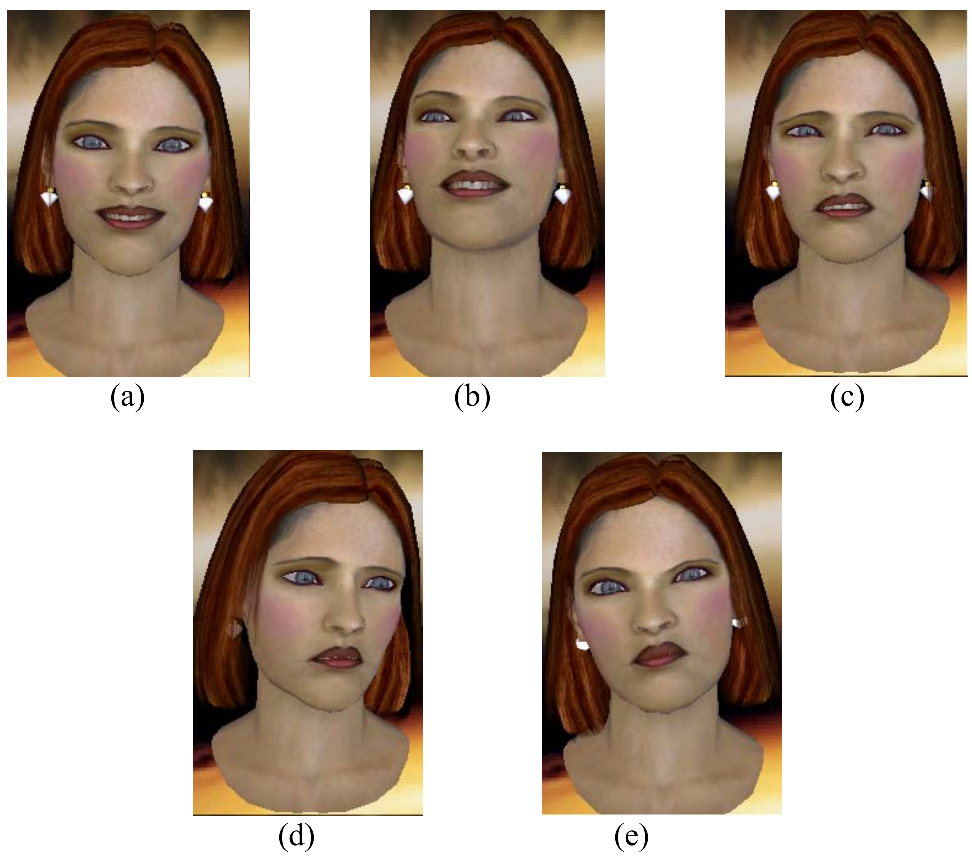


Figure 31: Five different facial expressions for the modeled emotions
(a) Happy, (b) Surprise, (c) Fear, (d) Sad, (e) Angry

4.8. Second Phase Development: Multi-Threaded Interface

To have Petra act and behave more naturally in social interaction, we implement our interface and ESG with several ARIA threads. This multi-threaded design, shown in Figure 32, enables her to navigate, receive the inputs through sensors, process the stimuli with the ESG and respond to them with a certain facial expression and the BSG *concurrently*.

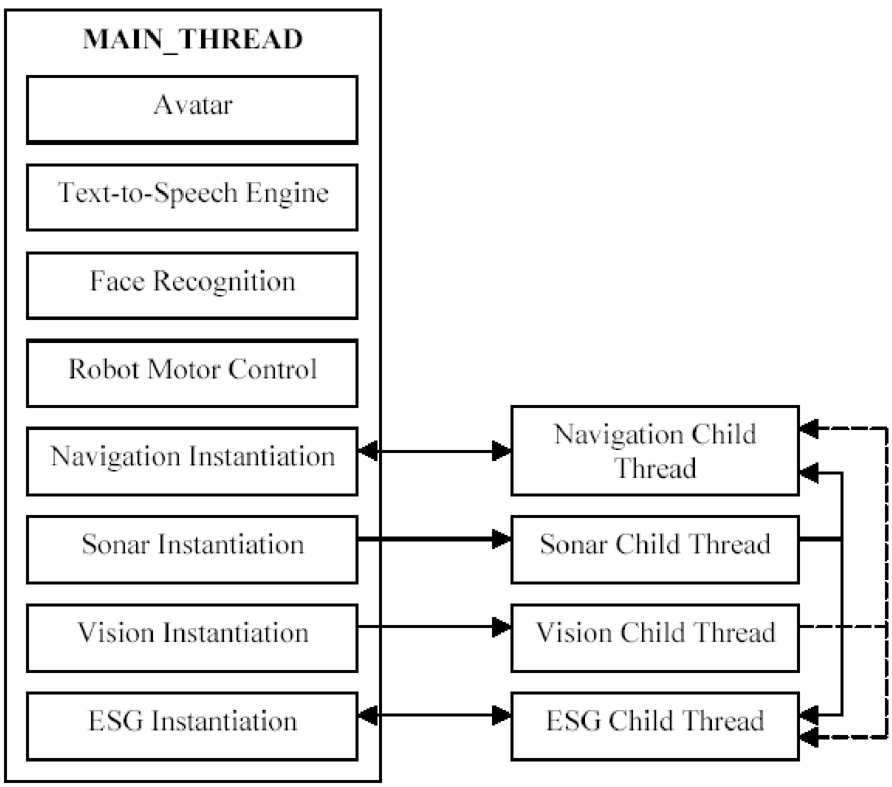


Figure 32: Our Multi-Thread Design

Initially, there is only one thread executed—the main thread. In this thread, we include the avatar, text-to-speech engine, face recognition system, robot motor control, navigation instantiation, sonar instantiation, vision instantiation (for obstacle avoidance), and ESG instantiation. Should she need to navigate when someone clicks the room-button, the child threads instantiated in the main thread—navigation, sonar, vision, and ESG—are executed. During navigation, the sonar and vision need to send the readings and images to both the navigation child thread, that is then passed back to the main thread to control the behavior of the robot motor system, and the ESG child thread, that calculates the emotion-like states and passes it forward to the main thread to change the facial expression in the avatar.

With all the components and advances described in this section—user-friendly interface, avatar, point-and-click map, human-height robot, multi-threaded interface, etc.— as well as the survey's result (Lisetti, Brown, Alvarez, and Marpaung, 2004), we believe that we are able to create an intelligent agent that can interact with humans more naturally. Thus, it enhances the Human-Robot Interaction (HRI).

CHAPTER FIVE: EXPERIMENTAL RESULTS

Data collected in the laboratory experiments confirm that Petra's internal state as well as the avatar and the behaviors, such as STAY_CENTER, AVOID_LEFT_WALL, AVOID_RIGHT_WALL, and WAIT, can be dynamically adapted with the changes in the environment.

In these experiments, for most cases, both sonar and vision are used for the navigation and obstacle avoidance system, but for certain conditions, only either one of them is used. We have to rely on one sensor for certain conditions such as when Petra (1) gets out from an office to navigate between the hallways, (2) almost hits either the left or right wall, (3) almost reaches the "L" or "T" junctions—the intersections between one hallway with the other.

Gets out from an office to navigate between the hallways: This case occurs when Petra either (1) finishes her tasks in a certain office suite (when she is in the room), or (2) waits outside the office and faces the door (when she never enters the room). In the first case—after finishing her tasks—she faces and heads to the door and moves forward to the center of the hallway. From this point, she continues her navigation to the next chosen room. In the center of the hallway, she either makes a left or right turn. But in the second case—after waiting outside the office for a period of time and deciding to move to another room—she decides whether to make a left or right turn and then moves to the center of the hallway and continues navigating. For both cases, the vision system needs to be turned off because it cannot detect the center of the hallway. Thus, Petra only relies on the sonar readings. The vision is turned back on when she is in the center of the hallway.

Almost hits either left or right wall: In this condition, Petra has to fully rely on the sonar because her vision system cannot measure her distance from the wall. The vision is turned back on when she is several hundreds millimeters away from one side of the wall.

Almost reaches the "L" or in the "T" junctions: When Petra almost reaches the L-junction, she has to turn off the vision system. Since she is too close to the wall in front of her and, unfortunately, there is no vanishing point in the captured image; she has to depend on the sonar information when making either a left or right turn to the other hallway. When she almost reaches the junction, she needs to detect the starting point of the other hallway (the coordinates of the right wall if she makes a right turn or the left wall if she makes a left turn) from a large difference between two consecutive sonar readings. From this point, she needs to move forward a certain fixed distance and make a turn. Afterward, she keeps moving forward until she detects another large difference between the two consecutive sonar readings (the sonar readings measured from the right wall if she makes a right turn and the ones measured from the left wall if she makes a left turn). The large difference indicates that she is already in the other hallway. The vision system is turned back on when she is in the center of the hallway, which is not too far after making a turn.

When reaching the T-junction, like the L-junction, the vision system is also turned off because she needs to detect the starting point of the other hallway from a large difference between the two consecutive sonar readings. When she finds the starting point of another wall (the right one if making a right turn and the left one if turning left), she needs to move forward to the center point of both hallways, make either a left or right turn, and then move forward until receiving another large difference between two consecutive sonar readings (the readings measured from the right wall if she makes a right turn and the ones measured from the left wall if she makes a left turn), which tells her that she is already in the other hallway. At this point forward, she needs to center herself between the walls, and when she is in the center position, she can turn the vision back on.

5.1. Experimental Method

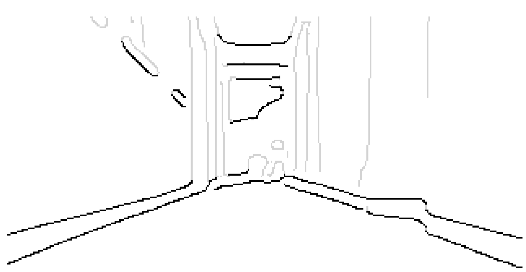
To demonstrate the claims of this thesis, a series of scenarios were executed. Scenario one demonstrates the situation in which the hallway has fewer obstacles during navigation and Petra finds the door, thus producing a happy schema. Scenario two demonstrates the situation in which the hallway has many obstacles, thus producing a sad schema to the memory for this particular movement. Scenario three demonstrates the situation in which Petra does not recognize someone, which leads to an anger schema.

5.1.1. Scenario one – navigation with fewer obstacles

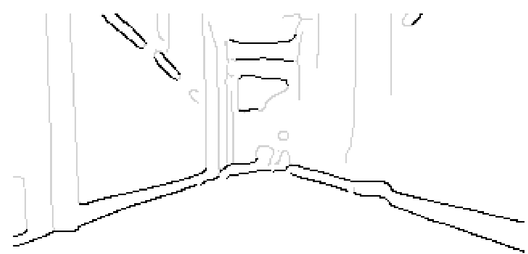
For the first scenario, Petra is asked to move from room 204, her default position, to room 211. Since both rooms are located in the same hallway, both sonar and vision are used concurrently to calculate the drifting rates and angle changes. The images in Figure 33, which are edged and smoothed, show that Petra encounters fewer obstacles during navigation.

(a)

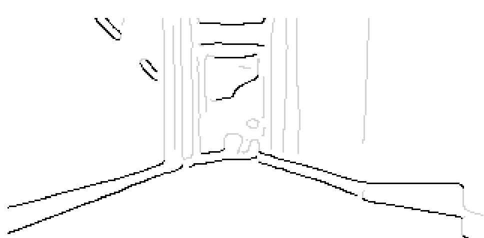
(b)



(c)

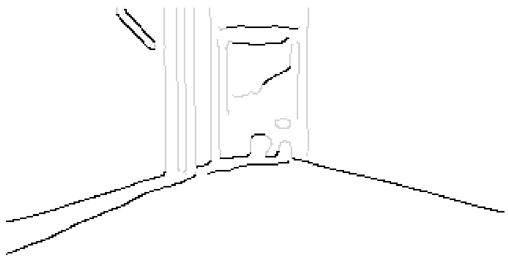


(d)

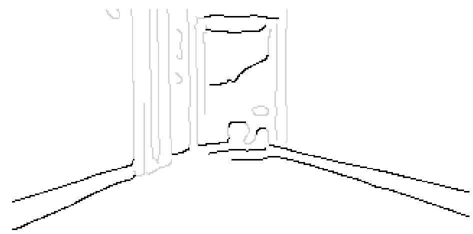


(e)

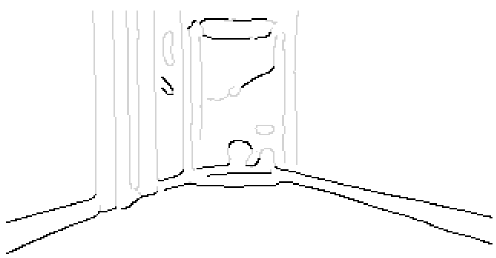
(f)



(g)



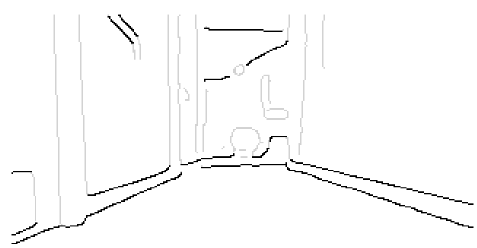
(h)



(i)

(j)

(k)



(l)

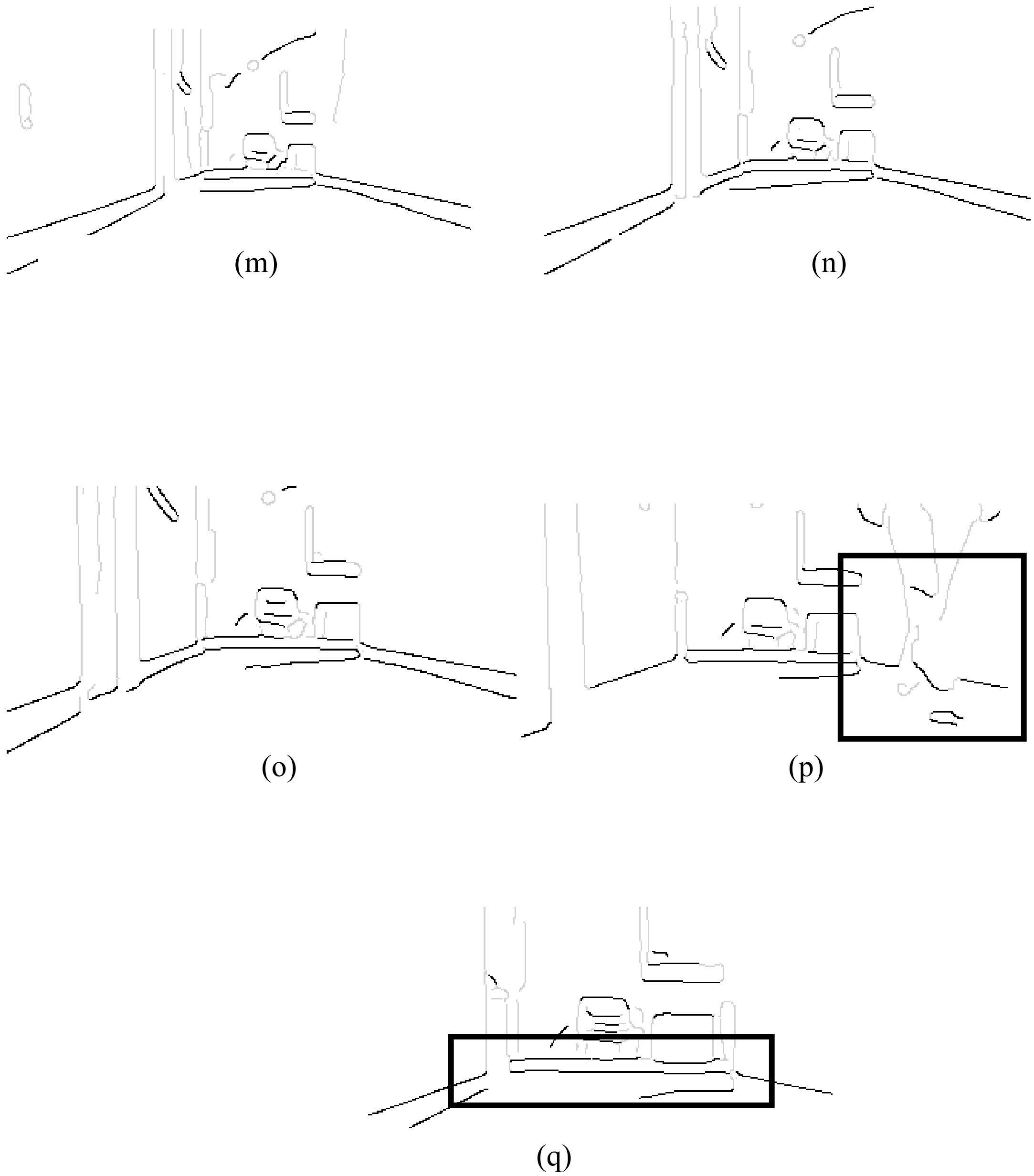


Figure 33: The edged and smoothed images for the movement from room 204 to 211

The only obstacle found is shown in Figure 33 (p) where Petra encounters a moving object. Thus, the moving object triggers the surprise emotion-like state for this cycle and she displays a surprise facial expressions. Petra can decide that an obstacle is a moving object from its absence in the image taken after detecting an obstacle (or Figure 33 (q)).

For most images in Figure 33 above, we can see that there is no obstacle encountered. The absence of the obstacle gives a happy emotion-like state for most cycles (except for the one in Figure 33 (p)). On the other hand, Petra also can find a door that leads to a happy state. Since Petra only encounters one obstacle and finds the door, this particular movement is associated with a happy state in her schema, shown in Table 6.

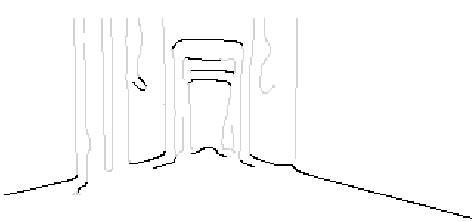
Table 6: A schema of moving from room 204 to 211

Components	**Values**
Emotion	Happy
Valence	Positive
Intensity	Medium
Focality	Event – from 204 to 211
Agency	Self
Modifiability	Medium
Action Tendency	Generalized readiness
Causal Chain	- Something good happened now - I wanted this - I do not want other things - Because of this, I feel good

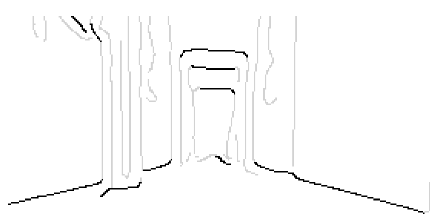
5.1.2. Scenario two – navigation with more obstacles

In the second scenario, Petra is asked to move from room 262 to room 252. Similar to the case in the first scenario, since the rooms are located in the same hallway, both sonar and vision are used concurrently to calculate (and correct for) the drifting rates and angle changes. The images in Figure 34 shows that during navigation, in several cycles, Petra encounters no obstacles in (a), (b), and (c); but on most cycles, Petra faces several obstacles in (d), (e), (f), and (g).

In initial execution from room 262, the hallway has no obstacles as shown in Figure (a)-(c). At these cycles, Petra is in a happy state. But several cycles later, as shown in Figure (d)-(g), the moving objects are detected, which triggers Petra to make several medium course corrections, triggering a sad state.



(a)



(b)

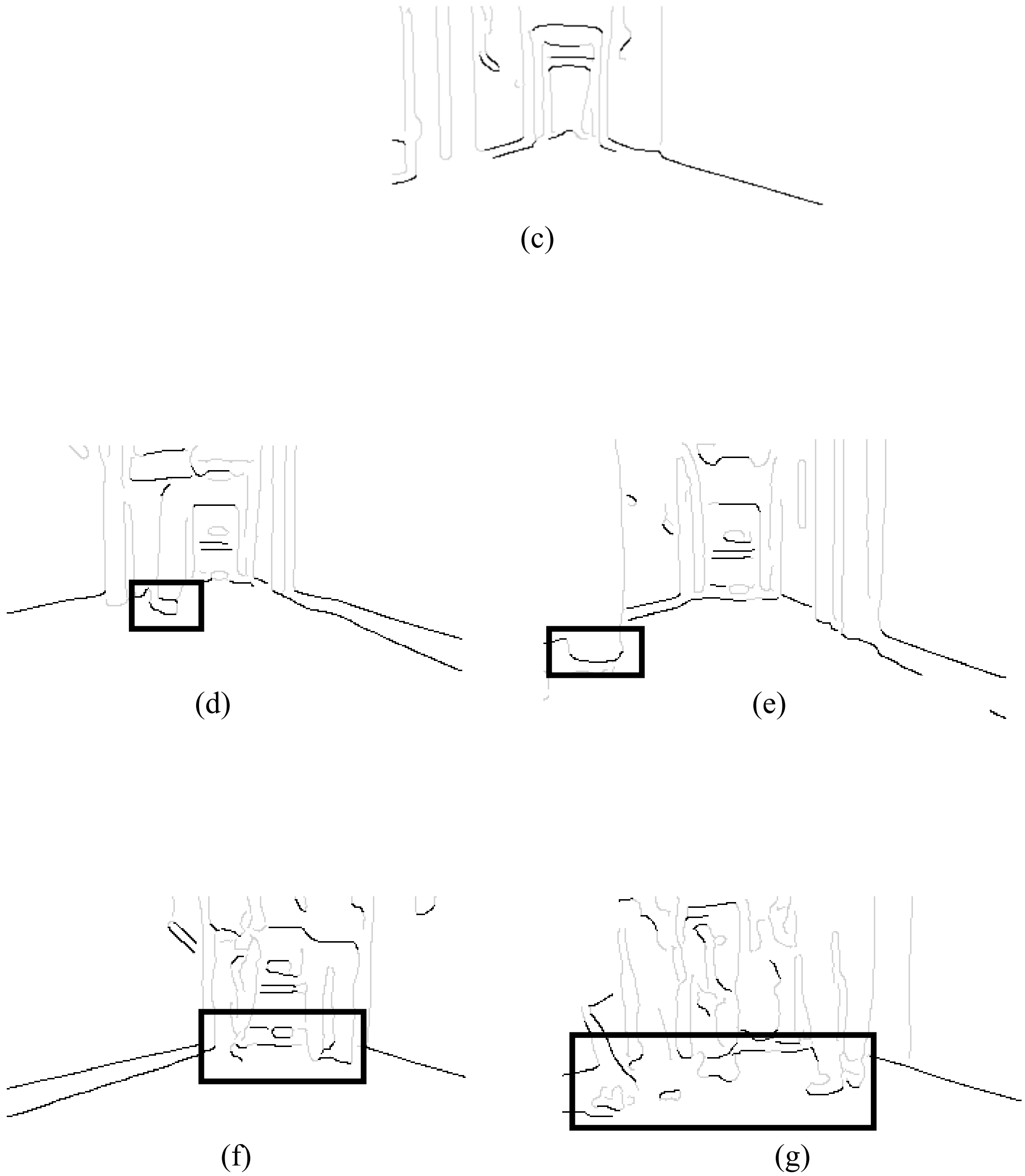


Figure 34: The edged and smoothen images for the movement from room 262 to 252

Because there are many obstacles and course corrections during navigation that reduces Petra's actual travel distances, by the end of the journey, she fails to detect the door. Her failure to detect one in addition to her previous emotional state—sad—caused by the obstacles triggers the creation of a sad schema, shown in Table 7, which is stored in her memory.

Table 7: A schema of moving from room 260 to 252

Components	**Values**
Emotion	Sad
Valence	Negative
Intensity	Medium
Focality	Object – from 260 to 252
Agency	Other
Modifiability	Medium
Action Tendency	Avoid
Causal Chain	- Something bad happen - I would want this did not happen - If I could, I would want to do something because of this - I cannot do anything

5.1.3. Scenario three – cannot recognize someone

Scenario three is the continuation of the previous one—scenario two. From scenario two, Petra is already in a sad state. As explained in "From Fuzzy Logic to Emotion-Like States" of the implementation section, the final emotion-like state is calculated based on the max ( ) function, which picks the highest value among the emotion-like parameters and makes it the final state. In our current situation, having a sad state implies that the sad emotion-like parameter is the highest among others.

Based on the fact that Petra fails to detect the door, she cannot get into the room in order to recognize someone. Due to her failure to recognize someone, her sad emotion state is changed to anger, and an anger schema is created and then stored in her memory.

Table 8: A schema of anger due to Petra's failure to recognize someone.

Components	**Values**
Emotion	Angry
Valence	Negative
Intensity	Medium
Focality	Object – Dr. So and so
Agency	Other
Modifiability	Medium
Action Tendency	Avoid
Causal Chain	- This person did something bad - I do not want this - Because of this, I want to do something - I would want to do something bad to this person

5.2. Summary

Data collected during laboratory testing scenarios demonstrate the changes in Petra's emotion-like state, which is dynamically adapted to the environment. For all emotion-like states generated, Petra is able to display appropriate facial expressions as well as to create schemata stored in the memory. In the future, these schemata can go through further learning and thinking processes to produce several other emotion-like states, the corresponding facial expressions as well as the appropriate behaviors.

With dynamically emotion-like state change as well as the facial expression, Petra is able to adapt her internal state with the changes in her external world. And with these facial expression changes, based on our survey (see Appendix A), the users can have better interactions with the robot that can express emotion and show facial expression.

CHAPTER SIX: FUTURE WORK AND CONCLUSION

The motivation for this research has been to propose the design of the sensory motor level as part of the three-layered architecture inspired by the Multilevel Process Theory of Emotion (Leventhal, 1979; Leventhal, 1980; Leventhal and Mosbach, 1983; Leventhal, 1984; Leventhal and Scherer, 1987) and implemented in the hybrid deliberative/reactive architecture robot. The change of facial expression, as the external communication corresponding to internal states, has been demonstrated. Data collected in Chapter 5 shows that our design works. The following sections will discuss the issues raised by this thesis.

6.1. Emotion versus traditional control

This thesis proposes the sensory motor level that serves as the first layer of the ESG. Both the schematic and conceptual levels are left for future work. The absence of both levels leave open issues of how the other factors, e.g., more behaviors, learning process, further and willful thinking, and etc. can be integrated in the ESG.

The simple implementation of this level raises the question of whether it would have been easier just to engineer the solution and do not use emotions to control the agent's behaviors. This thought can be true if we only think about the short-term goals of controlling its behaviors.

In the long run, the sensory motor level serves as the foundation for further work on ESG design. The schemata corresponding to the emotional experiences and the external inputs produced at this level do not go through further learning and thinking processes. Thus, they may lack some meaningful information from the agent's world. But the presence of these schemata in the memory can help the implementation of the other two levels where these schemata can be learned and thought further to produce meaningful and useful information than the one produced in the sensory motor level as well as the emotion-like states.

6.2. Enhancing Human-Robot Interaction (HRI)

In this thesis, we illustrate our robots' properties and functionalities, which aim at modeling emotions to enhance HRI—based on the results gathered from the survey (Lisetti, Brown, Alvarez, and Marpaung, 2004). After finding out that the idea of having emotions in a robot was warmly accepted, our next task was to find some ways to integrate emotions that could meet our needs.

There were many considerations taken in designing our user-friendly interface from both hardware and software sides. In short, for the hardware, we chose to model emotions in a human-height robot and displayed our interface in a touch screen. As explained in section 4.4, we believe that our shifting from Cherry to Petra has positively increased the social interaction. Instead of having an interface at a low level in Cherry, we can display it in a touch screen at a human height. Thus, the users do not have to kneel down to operate it and the touch screen eliminates the usage of keyboard and mouse because it can be easily operated with a stylus. For the software, we created our interface that fitted our needs. Based on Petra's functionalities as a tour guide and an office assistant, we choose to display the point-and-click map to ease the operation, choose the avatar to help people understand her internal state easier, choose the speech textbox to help people comprehend what she says in a noisy environment as well as to accommodate the hearing impaired. Thus, users can have better social interactions with the robot.

In this thesis, we do not claim that our design interface is the only way to enhance HRI. There are many creative things that can be done but we, as researchers as well as developers, need to think about our intended users so the HRI can still be further enhanced.

The other thing that we would like to have in our agents is personality. Having personality in an agent—e.g., humming while waiting, singing while navigating, etc.—can make it more human-like. Each personality can also be tailored to the agent's emotion transition. For example, a meek personality tends not to get angry easier compared to an aggressive one. If we can match the agent's personality with a user, the HRI can be enhanced, e.g., a meek person could interact with a meek robot instead of having that same person interact with an aggressive robot.

Besides having personality, we can also allow users to select different kinds of avatars, e.g., young Asian female, African-American male, old woman, Caucasian boy, Hispanic girl, etc., that can make the social interaction memorable. With these various options, users are able to pick the one that really attracts them. For example, many Asian males may choose a young Asian female over an old woman; an old male veteran may choose an old woman, etc.

We can also enhance HRI by tailoring the agent's conversation to users' interests. If the agent is able to communicate something of interest to the users, they may be able to interact, thus, enhancing HRI.

The descriptions above are only a few ideas that can be implemented to enhance HRI. The most important thing is to watch our intended users so that the developed ideas can be associated with their lifestyles as well as interests.

6.3. Scalability

In our proposed model, we have used several sensors, i.e., sonars, camera for face recognition, and camera for obstacle avoidance, to receive stimuli from outside world. The ESG then processes the stimuli with fuzzy logic. With these sensors, we can have a good navigation and obstacle avoidance system but it can be enhanced by having more sensors to capture more information. Thus, our architecture can be expanded to include several others, such as laser and thermal sensor, while processing the stimuli with fuzzy logic.

If more sensors are added, several other things, e.g., the PC's power, need to be considered. To enhance HRI, we have to ensure that the robot response time[23] is acceptable. If during running time it cannot respond in a short time, the users may be bored, thus decreasing their likeliness to interact with it and can deteriorate HRI.

[23] Response time is the difference between the starting time when the robot receives the stimuli and the ending time when the robot responses to these stimuli.

6.4. Summary and conclusions

This thesis has shown the design of the sensory motor level that serves as the basic foundation of the ESG. The accepted stimuli—sonar reading, vision navigation image, and face recognition image—need to go into the perceptual system. In this system, the stimuli are filtered from useless information in every cycle. The system produces some more meaningful information (but still fuzzy to certain degree), which becomes the inputs to the ESG. In the sensory motor level, the filtered inputs—valid sonar readings, vision navigation interpretation, and person's name—are processed further with TSK fuzzy logic that produces the emotion-like states as well as the appropriate and limited behaviors. At current implementation, the behaviors are triggered to avoid the robot's movement from hitting the walls and obstacles.

After calculating the emotion-like state, the inputs associated with each emotional experience are checked in Stimulus Evaluation Checks (SECs) where a schema, with several emotion components, is assigned to some appropriate values and is then stored in the memory for further work. As a final result, at every cycle, Petra is able to show her current internal state, which matches her current emotion-like state, through her anthropomorphic face, the avatar.

Although the implementation of this level is relatively simple, it can provide some basic information stored in the memory that can be used in future implementations of the schematic and conceptual levels. At these levels, this information can go through further learning and thinking processes to express many different emotion-like states as well as facial expressions. Thus, we hope that this work can pioneer the ESG implementation to enhance HRI, thus making HRI more like human-human interaction.

APPENDIX A: Publication I

C. Lisetti, S. Brown, K. Alvarez, and A. Marpaung. A Social Informatics Approach to Human-Robot Interaction with a Service Social Robot. *IEEE Systems, Man, and Cybernetics – Special Issue on Human Robot Interaction.* Vol 34. No. 2. May 2004

A Social Informatics Approach to Human-Robot Interaction with a Service Social Robot

Christine L. Lisetti, Sarah M. Brown, Kaye Alvarez, Andreas H. Marpaung

Abstract—**The development of an autonomous social robot, Cherry, is occurring in tandem with studies gaining potential user preferences, likes, dislikes, and perceptions of her features. Thus far, results have indicated that individuals (a) believe that service robots with emotion and personality capabilities would make them more acceptable in everyday roles in human life, (b) prefer that robots communicate via both human-like facial expressions, voice, and text-based media, (c) become more positive about the idea of service and social robots after exposure to the technology, and (d) find the appearance and facial features of Cherry pleasing. The results of these studies provide the basis for future research efforts, which are discussed.**

***Index Terms*— human-robot multimodal interaction, robot building tutorial, multimedia integration, emotion, personality, socially intelligent affective agents.**

I. INTRODUCTION

Increasing advances in the field of Artificial Intelligence (AI), AI robotics [1], behavior-based systems [2], [3], robot sensor fusion [4], [5], [6], robot vision [7], and robot emotion-based architectures [8], [9], [10], [11] have rendered feasible a variety of applications for human-robot interaction and collaboration. These include planetary exploration, urban search and rescue, military robotic forces, personal care and service robots (e.g., hospital assistance, home elderly care, robotic surgery), home appliances, entertainment robots, and more [12].

Although complete robot autonomy has not yet been accomplished, "the feasibility of integrating various robot entities into people's daily lives is coming much closer to reality. […R]obots now have the potential to serve not only as high-tech workhorses in scientific endeavors, but also as more personalized appliances and assistants for ordinary people" [12].

As robots begin to enter our everyday life, an important human-robot interaction issue becomes that of *social relations*. Because emotions have a crucial evolutionary functional aspect in social intelligence, without which complex intelligent systems with limited resources cannot function efficiently [13], [14] or maintain a satisfactory relationship with their environment [15], we focus our current contribution to the study of emotional social intelligence for robots.

Indeed, the recent emergence of affective computing combined with artificial intelligence [16] has made it possible to design computer systems that have "social expertise" in order to be more autonomous and to naturally bring the human – a principally social animal – into the loop of human-computer interaction.

In this article, *social expertise* is considered in terms of (1) internal motivational goal-based abilities and (2) external communicative behavior. Because of the important functional role that emotions play in human decision-making and in human-human communication, we propose a paradigm for modeling some of the functions of emotions in intelligent autonomous artificial agents to enhance both (a) robot autonomy and (b) human-robot interaction. To this end, we developed an autonomous service robot whose functionality has been designed so that it could socially interact with humans on a daily basis in the context of an office suite environment and studied and evaluated the design *in vivo*. The social robot is furthermore evaluated from a social informatics approach, using workplace ethnography to guide its design *while* it is being developed.

From our perspective, an interesting modeling issue therefore becomes that of *social relations*. In particular, we have chosen to focus our contribution to the field in addressing the technical goals of (1) *understanding how to embody affective social intelligence and (2) determining when embodied affective social intelligence is useful (or not).*

In order to determine answers to these issues, our approach is to develop a framework for computationally representing affective knowledge and expression based on cognitive modeling and to *concurrently* conduct surveys in order to investigate three areas: (a) human social intelligence, (b) robot social intelligence, and (c) human-robot social interaction.

a. Human social intelligence: One may ask whether the personality of the human affects how the human interacts with the robot. If so, how? Does it arouse specific emotions or behaviors? Which ones? In what contexts does this happen? Are these effects consistently observable, predictable, positive, or negative? Can we improve on these toward the positive? If so, how?

b. Robot social intelligence: Examples of such concerns are found in quests such as whether a

Manuscript received August 9, 2002, revised July 20, 2003. This work was supported in part by Office of Naval Research (ONR) under Grant N00014-03-1-0187.

Christine L. Lisetti, Sarah M. Brown, and Andreas H. Marpaung are with the Department of Computer Science, University of Central Florida, Orlando, FL 32816 USA (e-mail: lisetti@cs.ucf.edu, sbrown@cs.ucf.edu, marpaung@cs.ucf.edu).

Kaye Alvarez is with the Personnel Board of Jefferson County, Birmingham, AL 35203 USA (e-mail: kayealvarez@charter.net).

machine without emotions really is intelligent and autonomous. If not, how can emotions be modeled to increase robot autonomy? Can "no personality" in an intelligent agent (software or robot) be perceived by humans as a cold, insensitive, indifferent agent? If so, do these perceptions differ by specific groups of people, differentiated by age, gender, culture, sub-culture, etc.? Is it important to change the perceptions mentioned above in humans so that agents can be viewed as personable, helpful, even compassionate? If such is the case, can we identify the various contextual situations and applications when these agent properties might be beneficial, or even necessary? If emotions and personality are embodied in a robot, does it affect how people respond to it? If so, how and in what contexts? Should they resemble that of humans, or should they depart from them?

c. Human-robot social relationship: Finally, questions arise as to what kind of taxonomy of human-robot social "relationships" can be established, identifying numeric (e.g., one-to-one, one-to-many, many-to-many), special (e.g., remote, robo-immersion, inside), and authority (e.g., supervisor, peer, bystander) relationships [12] to determine what levels of "interpersonal skills" a robot would need in order to perform its role(s) effectively.

In Section 2, related research approaches are surveyed in terms of emotion modeling and emotion-based architectures as well anthropomorphic avatars and social informatics approaches to evaluate designs. In Section 3 the paradigm used for modeling emotional intelligence in artificial artifacts is set forth. Section 4 describes the actual implementation of mechanisms for endowing an autonomous mobile robot with affective social intelligence. In Section 5, the results of a survey conducted to evaluate the robot design and to determine exactly when embodied affective social intelligence is useful or not are produced. In addition, a discussion about the consequences of the study's results from a participatory perspective is provided. Finally, Section 6 discusses future research issues.

II. RELATED RESEARCH

A. Emotion-Based Robot Architectures

There have been several attempts to model emotions in software agents and robots and to use these models to enhance functionality. El-Nasr [17] uses a fuzzy logic model for simulating emotional behaviors in an animated environment. Contrary to our approach directed toward robots, her research is directed toward HCI and computer simulation.

Velasquez's work [10], [18] is concerned with autonomous agents, particularly robots in which control arises from emotional processing. This work describes an emotion-based control framework and focuses on affect programs which are implemented by integration of circuits from several systems that mediate perception, attention, motivation, emotion, behavior, and motor control. These range from simple reflex-like emotions, to facilitation of attention, to emotional learning. Although the approach is different, its motivation is similar to ours.

Breazeal's work [8], [9] also involves robot architectures with a motivational system that associates motivations with both drives and emotions. Emotions are implemented in a framework very similar to that of Velasquez's work but Breazeal's emphasis is on the function of emotions in social exchanges and learning with a human caretaker. Our approach is different from Breazeal's in that it is currently focused on both social exchanges and the use of emotions to control a single agent.

In Michaud's work [19], [20], emotions *per se* are not represented in the model, but emotion capability is achieved by incorporating it into the control architecture as a global background state. Our approach which chooses to represent the emotional system explicitly (as discussed later) differs from Michaud's in that respect. Although both Michaud and our approach revolve around the notion of emotion as monitoring progress toward goals, our work explicitly represents emotion and corresponds to a formal cognitive model.

Murphy and Lisetti's approach [11] uses the multilevel hierarchy of emotions where emotions both modify active behaviors at the sensory-motor level and change the set of active behaviors at the schematic level for a pair of cooperating heterogeneous robots with interdependent tasks. Our current approach builds on that work, setting the framework for more elaborate emotion representations while starting to implement simple ones and associating these with expressions (facial and spoken) in order to simultaneously evaluate human perceptions of such social robots so as to guide further design decisions.

B. Communicative Anthropomorphic Artificial Agents

Much research is currently underway on the subject of agent-based interaction [21], and agents of the future could promise to decrease human workloads and make the overall experience of human-computer interaction less stressful and more productive. Agents may assist by decreasing task complexity, bringing expertise to the user (in the form of expert critiquing, task completion, coordination), or simply providing a more natural environment in which to interact [22].

Specifically, there are a number of other related research projects that have studied the animation of computer characters/avatars in order to further the effectiveness of human-computer interaction [23], [24], [25], [26]. The current research aims at furthering progress in that area.

C. Social Informatics Approaches to Evaluating Human-Robot Interaction

Formally, social informatics is "the interdisciplinary study of the design, uses, and consequences of information technologies that take into account their interaction with institutional and cultural contexts" [27]. One key idea of social informatics research is that the "social context" of information technology development and use plays a significant role in influencing the ways that people use information and technologies.

As a consequence of these findings, we take a socio-technical orientation in order to understand the specific features and tradeoffs that will most appeal to the people most likely to use our system. We rely on a set of "discovery processes" for learning about preferences of people interacting with our robot, which include workplace ethnography [28]. Indeed, as made clear recently by the cognitive science community, people, the systems they use, and the interaction between the two, can no longer be studied and modeled in terms of isolated tasks and factual information, but rather in terms of activities and processes [29].

To date, few researchers use this technique in their research. Two instances were found in the literature. For example, a non-humanoid robot capable of human interaction and performing repetitive tasks is being used to test the feasibility of robots for aiding autistic patients in learning social interaction skills [30]. At Carnegie Mellon University (CMU), the importance of having an avatar and face tracking device on a social robot was tested using their robot, Vikia, by monitoring the length of interactions with the robot [23].

What is unique to our socio-technical approach is that we mix quantitative and qualitative research methods via survey research to guide our design and implementation *concurrently*. In other words, we use survey results from potential users to guide the design of our robots rather than completing our design and then gaining their feedback.

D. Personality Theory

Because of our socio-informatic approach, which is essentially to create robots that potential users will find both useful and pleasing, various individual difference factors are also of interest. In particular, does a person's age, sex, ethnicity, educational interests, or personality determine their reactions to service and social robots? Will one robot design satisfy all types of users?

The assumptions behind personality theories are that personality traits (a) are stable across time (i.e., moods and emotions are temporary states); (b) influence behavior, perceptions, and thought processes; and (c) can be inferred from behavior. However, theorists do not agree on the number of factors. For example, Eysenck [31] found three factors, Costa and McCrae [32] found five, 16 factors were found by Cattell, Eber, and Tatsuoka [33], Gough [34] found 18 factors, and Saville, Holdsworth, Nyfield, Cramp, and Mabey's [35] found 31 [36].

Nevertheless, there is one theory of personality that has become most prominent: Costa and McCrae's [32] five-factor model, also known as the *Big Five*. There are several reasons why the Big Five has become popular. First, over the years, several theorists have independently found five factors of personality (e.g., [37], [38], [39], [40], [41], [42], and [43]). Second, longitudinal and cross-sectional studies have found support for five factors. Third, five traits appear to emerge from other personality systems. For example, Krug and Johns [44] investigated Cattell et al.'s [33] 16 factors and found five underlying dimensions. Finally, five factor models are found to generalize across age, sex, and cultures [36].

The dimensions of the Big Five include *extroversion, neuroticism, openness* to experience, *agreeableness,* and *conscientiousness*. An *extrovert* is described as a person who is energetic, assertive, outgoing, social, excitement seeking, and who tends to experience positive emotions. A person who is *neurotic* frequently experiences anxiety, depression, and negative emotions. In addition, he or she is described as impulsive, vulnerable, and self-conscious. Individuals who are *open* to experience enjoy new experiences, are open to ideas and values, and are often described as persons who enjoy the arts (e.g., music, theatre, etc.). *Agreeableness* is characterized as a person who is trusting, altruistic, compliant, tender-minded, and modest. Finally, a *conscientious* individual is competent, dutiful, organized, achievement oriented, self-disciplined, and deliberate [36].

III. APPROACH TO EMBODYING AFFECTIVE SOCIAL INTELLIGENCE

A. Embodied Social Intelligence and Decision-Making

In order to understand when social relationships are needed in human-robot interaction or when the perception of such relationships need to be changed, social relations must be modeled. Emotions have a crucial evolutionary functional aspect in social intelligence without which complex intelligent systems with limited resources cannot function efficiently [13], [14], nor maintain a satisfactory relationship with their environment [15].

Emotions are carriers of important messages which enable an organism to maintain a satisfactory relationship with its environment. *Fear*, for example, serves the function of preparing an organism physiologically for a flight-or-fight response (blood flow increases to the limbs, attentional cues are restricted, etc.). *Anxiety*, on the other hand, serves the function of indicating that further preparation for the task at hand is needed.

Emotions greatly influence decision making (although sometimes dysfunctionally), more often than not for improved efficiency and flexibility toward a complex changing environment. Indeed, pure reasoning and logic have proven to be insufficient to account for true intelligence in real life situations. In the real world with all its unpredictable events for example, there is not always time to determine which action is best to choose, given an infinite number of possible ones and a set of premises.

Furthermore, different personalities will incline individuals to have different mental and emotional pattern tendencies. An agent with an *aggressive* personality, for example, will be predisposed to a *fight* response when experiencing fear, whereas one with a *meek* personality will be predisposed to *flee*. Predispositions, however, can be altered by conscious repression and/or adaptation.

B. The Multilevel Process Theory of Emotions

The multi-level process theory of emotions [45] diagrammed in Fig. 1 was chosen for our approach because it considers emotions as complex behavioral reactions to external events and internal thoughts and beliefs constructed from the activity of a hierarchical multi-component processing system which parallels nicely robot architectures (as explained later):

a. The *sensory-motor level* is activated automatically without deliberate planning by a variety of external stimuli and internal changes (e.g. hormonal levels). Affective reactions based on pure sensory-motor processes are reflex-like and are coarse-grained states as described in Section 3.3: information available at that level consists of valence and intensity (see Fig. 1 lower layer).

b. The *schematic level* integrates sensory-motor processes with prototypes or scripts of emotional situations having concrete schematic representations (see Fig. 1 middle layer).

c. The *conceptual level* is deliberative and involves reasoning over the past, projecting into the future, and comparing emotional schemata in order to avoid unsuccessful emotional situations (see Fig. 1 upper layer).

The multi-level process theory of emotions is particularly powerful for artificial intelligent design in that it enables various levels to be implemented, integrated, and tested incrementally. As exemplified with an emotion-based architecture for two cooperating robots [11], it furthermore matches closely hybrid/reactive deliberative architectures for robotic agents. Table 1 shows that relationship.

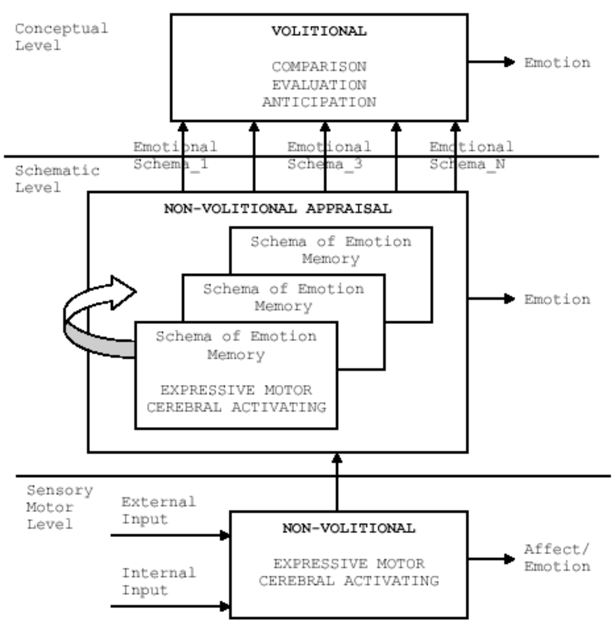


Fig. 1: Multi-level Process Affect/Emotion Generation

C. Affective Knowledge Representation (AKR)

In order to contribute to rendering artificial intelligent agents socially more competent, we combined and reconciled aspects of the main current theories of affect (e.g., [46]) and mood and emotion (e.g., [47], [48], [49]) into a simplified and comprehensive (but not complete) taxonomy of affect, mood, and emotion for computational Affective Knowledge Representation (AKR). The AKR is described in further details in [50].

1. Affect, Moods, Emotions, and Personality

We created the AKR in order to enable the design of a variety of artificial autonomous (i.e., self-motivated), socially competent agents in a variety of applications such as robotics [11], user-modeling [51], human-computer interaction [52], multi-agent

Table 1: Multi-level process of emotions vs. Hybrid reactive/deliberative [11]

Multi-Level Process	Hybrid Reactive/Deliberative
Conceptual • reasons about past and present emotions and projects into the future regarding possible consequences of action from anticipated emotion [66]	**Deliberative Planning** • reasons about past, present, future
Schematic • emotions control which behaviors are active through prototypical schemata • can be implemented with scripts [65]	**Assemblages of behaviors** • collections of behaviors are assembled into a prototypical schema or skill [3] • can be implemented with scripts [4]
Sensory-motor • emotions modify the motor outputs of active behavior	**Reactive behavioral** • active behaviors couple sensors and motor actions

systems, and distributed AI. The taxonomy of affective states is intended to differentiate among the variety of affective states by using values of well-defined componential attributes.

In short, in the taxonomy, each emotion is considered a collection of emotion components, such as its valence (the pleasant or unpleasant dimension), its intensity (mild, high, extreme), etc. The action tendency of each emotion [47] is also represented and corresponds to the signal that the emotional state experienced points to: a small and distinctive suite of action plans that has been (evolutionarily) selected as appropriate, (e.g. approach, avoid, reject, continue, change strategy, etc.).

Emotions are called "primary" or "basic" in the sense that they are considered to correspond to distinct and elementary forms of action tendencies. Each "discrete emotion" calls into readiness a small and distinctive suite of action plans that have been selected as appropriate when in the current emotional state. Thus, in broadly defined recurring circumstances that are relevant to goals, each emotion prompts both the individual and the group to act in a way that has been evolutionarily more successful than alternative kinds of prompting.

The number and choice of basic or primary emotions vary among different theories of emotion. We have selected the ones that seem to consistently reoccur across emotion theories. Their associated action tendencies are listed in Table 2.

Table 2: Action Tendencies

EMOTION	FUNCTION	ACTION TENDENCY
Fear	Protect	Avoid
Desire	Permit consummatory activity	Approach
Anger	Regain Control	Agnostic
Disgust	Protect	Reject
Anxiety	Caution	Prepare
Contentment	Recuperation	Inactivity

An emotional signal sent when a subgoal is achieved acts to prompt the individual to continue with the current direction of action. The signal sent when a goal is lost indicates a need to change the course of action or to disengage from the goal. Ensuing actions can be communicated to others in the same social group, which in turn, can have emotional consequences for the other individuals as well.

IV. SOCIAL SERVICE ROBOT IMPLEMENTATION

Enabling a computer for conversational interaction has been a vision since the creation of the first computers. While many components to a system capable of intelligent interaction with the user exist, having a believable agent capable of intelligent interaction is undoubtedly desirable. How can a *believable* emotional agent be created?

Part of the answer is to design agents whose behaviors and motivational states have some consistency. This necessitates (1) ensuring situationally and individually appropriate internal responses (in this case, emotions), (2) ensuring situationally and individually appropriate external responses (behaviors and behavioral inclinations), and (3) arranging for sensible coordination between internal and external responses [48].

Unless there is some consistency in an agent's emotional reactions and motivational states, as well as in the observable behaviors associated with such reactions and states, much of what the agent does will not make sense to the user.

Our robot, Cherry, currently has multiple internal states and external behaviors:

(1) maintaining and expressing a consistent personality throughout the series of interactions;
(2) experiencing different inner emotional-like states in terms of her progress toward her goals;
(3) choosing (or not) to express these inner states in an anthropomorphic manner so that humans can intuitively understand them;
(4) having an internal representation of her social status as well as the social status of her "bosses;" and
(5) adapting to the social status of the person she is interacting with by following acceptable social etiquette rules.

A. Hardware Overview

As an Amigobot from ActivMedia, Inc., Cherry's initial hardware included a Hitachi H8 processor, 1MB of flash memory, 2 reversible DC motors, 8 sonars, and a wireless modem. Her original functionality was limited to autonomous random wander movements or directed movements controlled by a stationary PC. As a result, many elements needed to be added to her hardware in order to increase her social interaction abilities. A small laptop was connected directly to the base of the robot to increase the programming capabilities, increase autonomy (i.e., the robot was no longer tied to a stationary computer), and allow the user interface to be displayed. Although we realize how impractical it is to have the interface at such a low level, it was not possible to create a platform at a higher level without causing her to tip over. Nevertheless, this design was implemented to begin our social robotic investigations, knowing that in the future we would be able to port the code to a different robot platform, as explained in "Future Research." To allow for face recognition and an eye-level vision for the operator, a FireWire camera was added to the top of an aluminum pole with a hub at its base. A detailed engineering tutorial on how she was modified is described in [53].

B. Robot Tasks and Functionality

In order to begin the inquiry on the modeling aspect of human-robot social relationships, we identified one specific application that appeared intuitively "social" enough to start generating interesting, relevant results.

Cherry was designed and programmed to participate in a number of office activities and to play a variety of social roles within an office suite. The algorithms designed for Cherry's roles include:

(1) her master's favorite office gopher: a 1-to-1 master-helper human-robot relationship;
(2) her department members' favorite gopher: a many-to-1 masters-helper human-robot relationship; and
(3) her department tour guide for visitor(s): another many-to-1 human-robot relationship.

Master(s)-Centered Gopher: Another important task Cherry can perform is delivering documents or bringing soda cans, which are deposited in her delivery cup, to a specific professor or staff member. A copy of the Computer Science map was created on Cherry's laptop interface to enable users (for now only one user at a time) to point and click to the location on the map he or she wants Cherry to go. Menu options are also available to choose a specific professor's office by last name. This feature will be described in more detail below.

Tour Guide Information for Faculty Offices and Faculty Research Interests: Another task Cherry can perform is to give meaningful and instructive tours of the faculty offices. In order to give Cherry knowledge of who works where so that she could introduce each researcher, each office on the map was linked with each professor or staff's facial image and current research interests (available from our UCF Computer Science web site and integrated in Cherry's software). In this way, Cherry has the capacity to introduce someone once she reaches his or her office.

C. Building Office Suite Map

ActivMedia Mapper [53] software was used to create a map of our Computer Science office suite in order to have the ability to create (1) a simple point-and-click navigation system and (2) a built-in grid system used in the navigational portion of the interface.

The robot is able to use its sonars to navigate around small and moving objects. As a result, only walls and large permanent obstacles needed to be drawn into the map. The robot's vision system for collision avoidance will be described later as future research.

The map associates the layout of the office suite and each office's corresponding suite number. It also includes information relating the name of each professor and staff member to their corresponding office numbers. In this way, the user can point and click on the office in order to dispatch Cherry to the office desired.

The map therefore provides quick and simple direction for Cherry. Because the map is completely accurate, it also provides the basis for the (x,y) coordinate system.

D. Eye-Level Vision and Face Recognition

The robot interface was also integrated with Identix face recognition code [54]. Cherry has the ability to take pictures of people she encounters with her eye-level camera, and to match them to her internal database of photographs of faculty, staff, and students who work in the Computer Science building.

E. Social Status and Greeting

Not only does face recognition abilities enable Cherry to recognize who she encounters, but also to greet different people according to their university status. These social status codes enable her to know what greeting is socially acceptable. In general these are clearly context and/or culture-dependent.

In the current case, they are limited to the distinction of social status within the UCF Computer Science Department: a Full Professor is greeted with more deference than a Graduate Student, by associating the title of "Professor" at the beginning of the greeting, versus addressing the person by their first name if the person is recognized as a graduate student, or yet by preceding the last name with Ms. or Mr. if the person is a staff member.

F. The Avatar

The avatar created is arguably the most important aspect of the robot interface. Indeed, with new advances in graphics over the past couple of years, artificial graphical representation of animated anthropomorphic faces have become realistic enough to convey subtle facial expression changes, skin tone, etc. Given how humans have developed over century of evolution a very efficient system to perceive and interpret facial expressions in human-human communication exchanges, the current approach aims at developing a scheme for human-robot interaction that exploits the natural human capacities to understand the meaning of facial expressions as they relate to internal state.

Cherry's face, shown in Fig. 2, was created using Haptek's People Putty [55] and was designed to be a 20-something year-old young woman who is both attractive and able to believably demonstrate being upset or angry. The avatar was designed to mimic human movement by incorporating random head and eye movements as well as lip movements as she spoke.

In order to facilitate Cherry's social interactions with humans, the avatar is present on the laptop (e.g., Cherry's user interface) and has voice capabilities, which allow her to speak to the user in natural language. As mentioned before, as a tour guide, her current tasks are to explain a variety of facts: who she is, what her mission is (namely the UCF computer science tour guide), which professor works in what office, what a particular professor is researching, what a professor's office hours are, and so on.

G. Speech and Voice

Haptek not only provides the means to create an avatar, but also to equip a robot with an appropriate voice. Selections include various male, female, and robotic voices, including voice simulations in space, in a stadium, on a telephone, and whispering. Because we wanted the avatar to be as human-like as possible, we decided to incorporate the standard female voice.

Fig. 2: Cherry's Neutral Facial Expression

H. Facial Expressions for Effective Communication

As surveyed in Lisetti and Schiano [56], since Darwin [57], the central preoccupation of researchers interested in the face has been to correlate movements of the face primarily with *expressions of inner emotional states*. The advocates of this view, the "Emotion View," are not all homogeneous in their opinions, but they do share the conviction that emotions are central in explaining facial movements [58], [59].

The "Behavioral Ecology View," on the contrary, derives from accounts of the evolution of signaling behavior, and does not treat facial displays as expressions of emotions, but rather as *social signals of intent,* which have meaning only in social contexts [60], [61]. [24]

These observations motivated the inclusion of facial expressions in our interface, with the intuition that humans would relate to and understand better a robot with an anthropomorphic face able to express internal states in a manner consistent with the one naturally used and understood by humans.

Currently, Cherry can display different facial expressions with different intensities, which, as explained later, correspond to her different inner states: neutral, frustrated, sad, and angry, as shown in Fig. 3 (a-d):

Fig. 3: a. (Left) Neutral Facial Expression

b. (Right) Frustrated Facial Expression

[24] More recently, facial expression has also been considered as an *emotional activator* – i.e. as a trigger – contrary to being viewed solely as a response to emotional arousal [62], [63], [64].

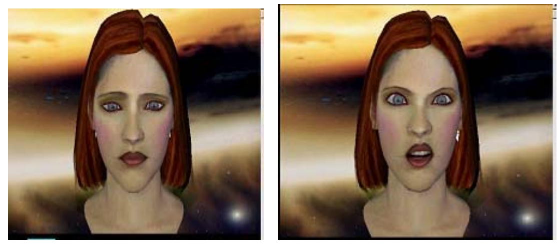


Fig. 3: c. (Left) Sad Facial Expression

d. (Right) Angry Facial Expression

I. Expression of Culturally-Independent Semantic Descriptions of Emotion Concepts

In order to enable our robot to express its internal emotional states in natural language as well, we adapted the semantic meta-definitions of emotion concepts using a limited set of language-independent primitives developed by Wierzbicka [49]. The semantic meta-definitions have the advantage of being culture-independent as they describe the causal chain that led to that emotion. A causal chain of events describes the subjective cognitive experience components that are associated with the emotion, the beliefs, the goals, and the achievement (or lack of) of those goals. These components are associated with each emotion and are spoken via speech synthesis so that the agent can *verbally* express and describe the cognitive interpretation of its state. For example, the causal chain for *frustration* is "I want to do something, I cannot do it, and because of this, I feel bad". More examples can be found in [65] again derived from Wierzbicka's work [49], and although slightly unnatural, we chose to use them in order to avoic ethnocentric language for our artificial agent. Furthermore, we also want to later be able to easily complete the uttered sentences with the actual objects of emotions, goals etc., and replace primitives like "something" (as above) with the actual object of frustration. For example, the robot will be able to identify the "something" that it is unable to accomplish in the *focality* of the causal chain. It will then say "I am frustrated because I want to deliver a message to Dr. So-and-so, and I cannot do it; because of this, I feel bad."

J. Internal States

Both a bottom-up and a top-down approach were adopted to design Cherry's architecture. She has the beginning of some social expertise in terms of associating a variety of *external expressive behaviors* with her various *inner states:*

(1) Frustration: Cherry reaches a state of frustration when she finds that an office to which she was send to has a closed door, or she cannot recognize the faculty or staff member inside the office. She expresses her internal frustration with the facial expression shown in Fig. 3b and with speech "I want to do something, I can't do this, because of this I feel bad."

(2) Anger: Cherry reaches an angry state when, after waiting for a long time, an office door still remains closed, and the action tendency activated will "motivate" her to change her current relationship with the environment and regain control. Anger is expressed with facial expression (Fig. 3d) and with speech "Something bad happened, I don't want this, because of this, I want to do something, I would want to do something bad to this object".

(3) Discouragement: Cherry reaches a discouraged state when, after waiting for a while, an office door still remains closed. She expresses sadness with the expression shown in Fig. 3c and with the speech "Something bad happened, I would want this did not happen, if I could I would want to do something, because of this I can't do anything."

The initial choice of specific internal states for Cherry was, on one hand, motivated by a desire to test how her different behavior affect real people behavior and their reaction to her (depending on their own personality, age, gender etc.), and on the other hand, to later be able to study the design of artificial agents in collaborative human-robot group settings.

These inner states – dynamically measured in terms of her current relationship with her environment and goals – will need to be integrated with the external behavior for a consistent system [48]. Currently, each level functions separately.

For the current application, the robot action tendencies (AT) associated with its emotion are related to its tasks and shown in Table 3.

Table 3: Cherry's Action Tendencies

EMOTION	AT for Cherry	ACTION TENDENCY
Happy	Guide/Deliver	FreeActivate
Neutral	Guide/Deliver	ContinueNormalActivity
Frustrated	ReturntoMaster	ChangeCurrentStrategy
Angry	RemoveObstacle	RegainControl
Discouraged	GiveUpTask	ReleaseExpectations

K. Emotion Dynamics

1. External Events as Inputs

Transitions among the various emotional states are caused by environmental inputs or responses to the system, and they are divided into categories of positive progress toward goals and negative progress toward goals. Using this dynamic model, we can predict that an agent that is in a HAPPY state will remain HAPPY given positive inputs and could become FRUSTRATED given a series of negative inputs towards its goal (e.g., obstacles of some sort depending on the context).

Currently, Cherry has a limited number of states to transition to and from: happy, neutral, frustration, discouragement, and anger as shown in Fig. 4.

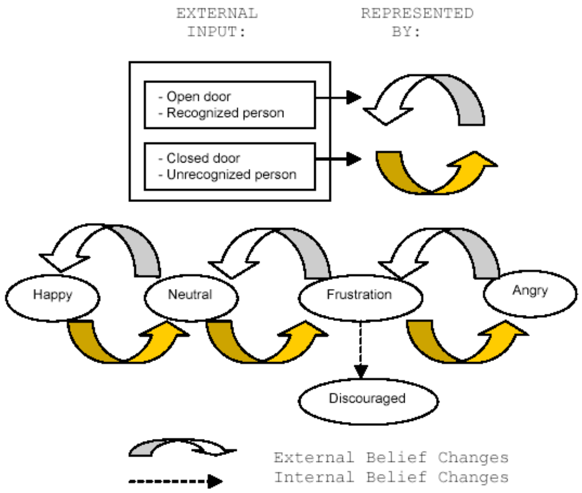


Fig. 4: Transitions between Emotional States

Transitions are based on negative inputs from the environment in terms of her success in (1) finding the door to the office that she was sent to open and (2) in recognizing someone in that office.

2. Internal Beliefs as Inputs

An individual's emotions can change in reaction to an event, and these changes may also be the result of their own efforts, not simply the result of an independent process directed by external events or social rules. Emotional changes indeed occur as a result of a number of processes.

A simple example is one where a negative internal belief regarding the subjective perception of *modifiability* of the current situation such as "I can't do this" keeps the agent in its current DISCOURAGED state forever. Should the agent manage to change its internal belief to a positive input in the form of an enabling belief (e.g., "I can indeed do this"), the agent would switch to a HOPEFUL state. Other examples of such internal self-adjustments abound [66].

These mental modal beliefs described in [50] are part of an affective knowledge representation scheme, which enables such transitions to occur. Currently, Cherry's internal beliefs such as *modifiability*, *certainty*, and *controllability* are not active in this version of implementation. Furthermore, depending upon the programmed personality traits, the agent can experience various tendencies toward specific sets of emotions.

L. Web-based Command-and-Control

To allow users the ability to control Cherry from their desktops (rather than having to stoop toward the floor to manipulate Cherry's laptop), the laptop was connected to the university network via a wireless Ethernet card.[25]

M. Cherry's Web-based eye-view of the world

Because a robot may take a "wrong turn" or intrude upon someone unintentionally, a vision aspect was integrated into the user interface. Not only is the image of what the robot can "see" (with the camera at eye-level) displayed on the user interface, but the image can be broadcasted via the Web to allow multiple users to view her actions at once.

This aspect of the complete user interface is partly for user interest, but mostly to prevent the robot from failing to reach an intended goal or advancing to an unsafe region, such as a stairway, due to inaccurate navigational systems during the testing process.

Using TeVeo webcam video streaming software, images can be broadcasted from Cherry's camera to the Web. Cherry's eye-level camera, and potentially another camera mounted nearer to her base, can provide a Cherry's-eye-view" of the world to users via access to the Web.

[25] We are searching for better ways to display the web interface in order to (1) reduce potential interferences and (2) get a better refresh rate and color display than WinVNC can provide. The subtle coloration and frequent subtle facial movements of our avatar caused by WinVNC will be described later.

N. The Complete Integrated Robot

Cherry's interface was written in Visual C++ and incorporates the avatar, speech, video, face recognition, and navigational map elements. We believe that the layout and simplicity of use will make the robot more accepted as a service robot and provide an easy and enjoyable way for people to interact with her. The avatar, map, eye-level vision, and menu options can all be seen in the integrated user interface in Fig. 5.

Finally, to create a non-intimidating genre of technology, and to give her an aesthetically pleasing appearance acceptable for a home, Cherry was dressed with feathers. This also has the advantage of avoiding issues such as raising user's expectations about her current abilities and limited intelligence.

V. DESIGN EVALUATION FOR SOCIAL ROBOT

Taking a social informatics co-evolutionary approach to the study and design of technology and social structures, this bi-directional approach enables us to start testing and evaluating the interface with human subjects *while* Cherry's functionality is being designed. We believe this approach helps to ensure maximum success in her functionality, interface design, and acceptance.

A. Study One: Preliminary Investigation

The first study was a preliminary investigation to determine whether our robots' features needed to be adjusted. Specifically, the objectives of the first study were to assess (a) whether Cherry's avatar and voice features were acceptable, (b) whether the avatar of a second robot under development, Lola, was acceptable, (c) opinions towards service robots, and (d) opinions towards robots with personality and emotion capabilities.

Method

Sample: The sample included 25 students and staff members from the engineering and computer science departments. There were 8 females and 17 males: 1 Hispanic, 16 Caucasians, 6 Asians, and 2 Native Americans. Their ages ranged from 18 to 55; however a mean age could not be calculated because the question asked the participants to specify their age range (i.e., 18-25 [n = 19], 26-35 [n = 2], 36-45 [n = 2], 46-55 [n = 2], and 56+ [n = 0]).

Procedure: The participants were given a demonstration of Cherry's features and social capabilities and were shown the avatar developed for Lola. The subjects then completed a questionnaire regarding their reactions to Lola's avatar and Cherry's features and appearance. In addition, the questionnaire also asked for their opinions of service and social robots.

Questionnaire: The questionnaire included 38 items: 4 demographic items (i.e., status, sex, age, ethnicity); 15 items assessing personality characteristics; 4 open-response items; and 15 items assessing their reactions to Lola's avatar, Cherry's appearance and features, their opinions of robots with personality and emotion capabilities, and their opinions of service robots in general. The personality items were not used in the analysis due to the sample size not being conducive for confirming the reliability and factor structure of the scale. In addition, the 4 open-response items were not used in the analysis, as a coding technique to enter the data into SPSS was not created. The purpose of these items was to determine why individuals liked or disliked Cherry's avatar and voice, Lola's avatar, and the idea of a robot with a personality.

The remaining 15 items included: two items regarding Cherry's avatar, 3 items referring to Cherry's voice, 1 item with regards to Lola's avatar, 6 items referring to opinions of robots with emotion and personality capabilities, and 3 items regarding opinions of service robot features. Two 5-point response options (i.e., 1 = definitely/extremely, 5 = not at all) were used with all but one item. The item, Which communication method would you prefer a robot use to inform you about the difficulties *it is having while accomplishing tasks?,* had three response options: human-like facial expressions of frustration, text-based list of commands the robot could not execute, or both.

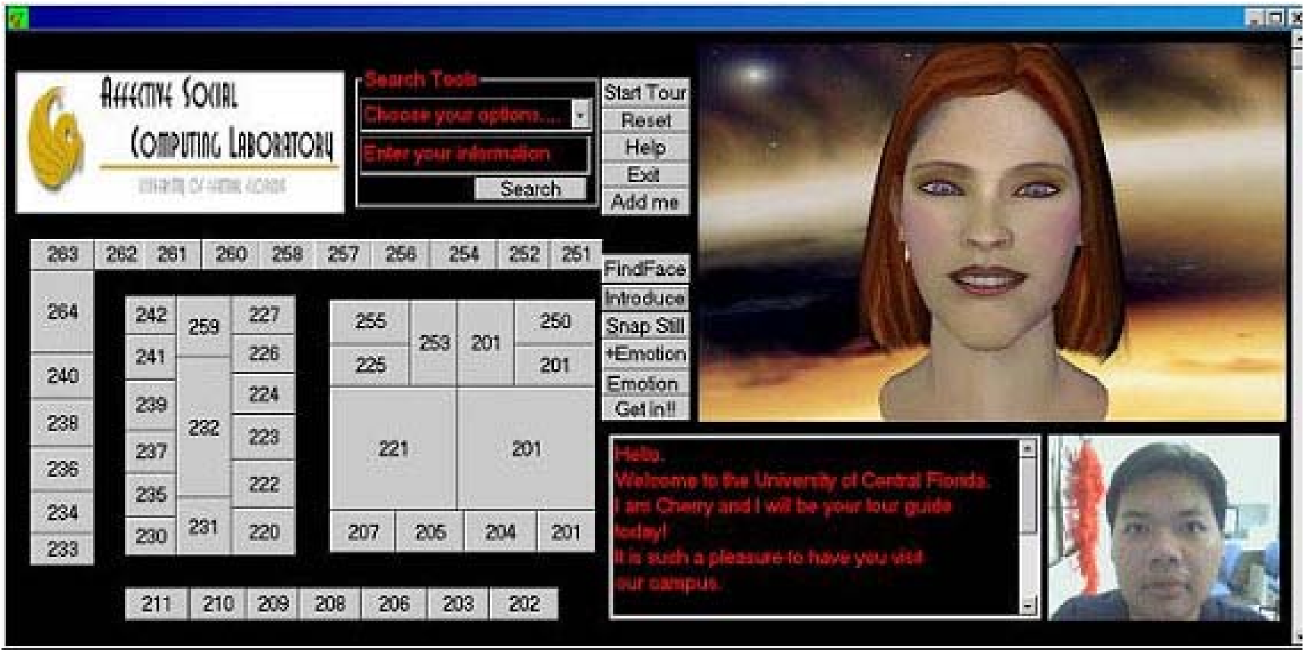


Fig. 5: Cherry's Complete Integrated Interface

Fig. 6 Cherry equipped for Social Interaction

Results: The average responses to the items regarding the two avatars were investigated first. The results revealed that, overall, the participants liked Cherry's avatar (*M* = 1.96, *SD* = .73) and did not like Lola's avatar (*M* = 3.43, *SD* = 1.16). In addition, overall, the participants enjoyed interacting with a robot having a human face (*M* = 2.38, *SD* = 1.01). The three items regarding Cherry's voice were summed and averaged. The average response to her voice (*M* = 2.53, *SD* = .99) indicated that the participants were pleased with the robot's voice and did not feel that her avatar mismatched her voice.

Overall, the participants felt that a robot with personality and emotion capabilities was a good idea (*M* = 2.10, *SD* = .99). In addition, they felt that a robot displaying positive emotions was acceptable (*M* = 1.56, *SD* = .92); however, they did not particularly like or dislike the idea of a robot displaying negative emotions (*M* = 3.00, *SD* = 1.44) or displaying frustration with people (*M* = 3.20, *SD* = 1.47) and objects (*M* = 2.96, *SD* = 1.49) interfering with its tasks.

With regards to service robots, the participants indicated that they liked the idea of a robot serving as a tour guide (*M* = 1.91, *SD* = 1.31) and a gopher (*M* = 1.48, *SD* = .81). Finally, on average, the participants preferred that a robot communicate its difficulties completing a task with both a human-like expression of frustration and a text-based list of commands it could not execute (M = 2.44, SD = .87).

B. Study Two: In-depth Investigation

Once determining that Cherry's avatar and that service and social robots were acceptable to people, a second, more extensive study was planned. The questionnaire items were revised to include more items regarding Cherry's overall appearance and specific features. In addition, more items regarding attitudes towards social and service robots were developed. Of particular interest was whether a person's demographic characteristics determined their responses. Therefore, the item regarding the age of the participants was changed to gain their actual ages and items asking for their major and department were added. Although it was not possible to determine if educational interests were related to responses in this study, we added these items for future investigations. The degree of experience individuals have interacting with or working on robots may also influence their reactions to robots; therefore two items regarding experience with robots were also added. Finally, in order to determine whether an online demonstration of reactions to Cherry would be feasible (potentially useful for future tele-medicine patient assistance and monitoring), items regarding how comfortable individuals would be with a robot broadcasting images to the Web were created.

The objectives of this study were to determine whether (a) the survey we created meets psychometric standards; (b) perceptions of and reactions to service robots, social robots, and Cherry differ by age, sex, ethnicity, or personality; (c) exposure to Cherry changed perceptions of service robots and/or social robots; (d) the features and appearance of Cherry were acceptable; and (e) individuals would be comfortable with a robot taking their picture and broadcasting images to the Web. The personality questionnaire developed for the current study is based on the Big Five theory of personality described in the "Related Research" Section.

Sample. The sample included 56 undergraduate students enrolled in a psychology course. There were 42 females and 14 males: 5 African Americans, 7 Hispanics, 34 Caucasians, 4 Asians, 5 individuals indicating mixed ethnicity, and 1 subject who did not report their ethnicity. Their ages ranged from 19 to 33 with a mean of 23.04 years (SD = 3.11).

Procedure. The participants completed a pre-questionnaire, which included items regarding their demographics, their opinions about service robots, and their opinion of robots with personality and emotion capabilities. After completing the pre-questionnaire, Cherry's features were described and a demonstration of her capabilities was presented. The subjects then completed a post-questionnaire regarding their reactions to Cherry's features and appearance. In addition, in order to determine whether exposure to Cherry changed their opinions regarding robots, the post-questionnaire also asked for their opinions of service robots and robots with social capabilities.

Pre-Questionnaire. The pre-questionnaire included 21 items: 6 demographic items (i.e., sex, age, ethnicity, major, department) and 15 items regarding their experience with robots, their opinions of service robots, and their opinions of robots with a personality and emotion capabilities. A 5-point Likert-type scale was used for 14 of the 15 items. The remaining item, *Which communication method would you prefer a robot use to inform you about the difficulties it is having while accomplishing tasks?,* had 3 responses to choose from: human-like expressions, text-based list of commands it could not execute, or both. Two items determined the participants' experience with robots (i.e., *How often do you interact with robots?* 1 = daily, 5 = none, and *What level of experience do you have working with or on robots?* 1 = high, 5 = none).

Five items assessed their opinions of service robots in general. The 5-point response options were of two types. For example, the item *Do you feel robots can be useful outside of an industrial setting (e.g., factories)?* included the following response options: 1 = definitely, 2 = pretty much, 3 = somewhat, 4 = a little, and 5 = not at all. The item, *How comfortable would you be with a robot serving as an assistant to help you remember appointments, grocery lists, etc.?*) included the response options of: 1 = extremely, 2 = very, 3 = moderately, 4 = somewhat, and 5 = not at all.

An additional 5 items asked participants about their opinions of robots with personality and emotion capabilities. For example, *Do you think giving a* *robot a personality is a good feature?* and *Do you feel that interactive robots should display emotions, positive or negative*? (1 = definitely, 5 = not at all). The final three items of the survey asked participants how they would feel about a robot taking their picture and having the images broadcasted on the Web.

Post-Questionnaire. The post-questionnaire included 38 items: 15 items assessing personality characteristics based on the Big Five personality theory and 23 items assessing their reactions to Cherry's appearance and features, their opinions of robots with personality and emotion capabilities, and their opinions of service robots in general. Three items for each of the five personality characteristics were developed (i.e., *I am sometimes shy and inhibited; I easily get nervous; I usually cooperate with others; Most often, I do a thorough job;* and *I enjoy art, music, and/or literature*).

Eight items assessed the subjects' reactions to Cherry's appearance, features, and social capabilities. The same two 5-point response options mentioned above were used. For example, *Did you enjoy interacting with a robot that has a human face?* had the 1 = *extremely* to 5 = *not at all* response options. The item, *Do you think the text box feature is helpful for understanding what Cherry says?* included the 1 = *definitely* to 5 = *not at all* scale. Six items assessed their opinions of service robots in general. The item *Which communication method would you prefer a robot use to inform you about the difficulties it is having while accomplishing tasks?,* was repeated in the post-questionnaire in order to determine if exposure to Cherry changed their preference for communication method. Other items included questions such as *Would you prefer a robot without a human face?* and *Would you like a robot to give you a tour of a building?* (1 = definitely, 5 = not at all).

An additional 8 items asked participants about their opinions of robots with a personality and emotion capabilities. In order to determine whether exposure to Cherry changed their opinions regarding social robots, two items from the pre-questionnaire were repeated in the post-questionnaire: *Do you think a robot with a personality is a good feature?* and *Do you think that having a robot display emotions could make them more accepted into everyday roles in human life?* (1 = definitely, 5 = not at all). Two additional items from the pre-questionnaire were also repeated; however, they were assessed with two separate items each. For example, the item *Do you feel that interactive robots should display emotions, positive or negative?* was assessed with the items: *Do you feel that interactive robots should display positive emotions, such as happiness and surprise?* and *Do you feel that interactive robots should display negative emotions, such as discouragement, frustration, and anger?* (1 = definitely, 5 = not at all).

The pre-questionnaire item, *Do you feel it would be appropriate for a robot to get angry or upset with an obstacle or person that interferes with a robot's task?* was measured with the items *Do you think it would be appropriate for a robot to communicate frustration or anger towards a person that interferes with its task?* and *Do you think it would be appropriate for a robot to communicate frustration or anger toward obstacles (i.e., walls, boxes) that interfere with its task?* (1 = definitely, 5 = not at all). The final item of the post-survey asked participants how important a person's overall appearance is to them when interacting with him or her. This question was asked in order to determine whether Cherry's physical appearance might hinder interactions with her.

Analyses. Five statistical analyses were performed with the data. Reliability theory suggests that any measurement technique, particularly in the behavioral sciences, contains some degree of error. The more error a test contains, the less reliable the results. Therefore, estimates of reliability are important to calculate before any other analyses are performed. Reliability estimates range from zero to one: the larger the number, the more reliable the test. Estimates equal to or greater than r = .80 are recommended when the goal is to make comparisons between groups [67]. The reliability estimates for the items measuring attitudes towards service robots from the pre- and post-questionnaires were r = .85 and r = .51, respectively. For the items assessing attitudes towards social robots (e.g., with emotion and personality capabilities) in the pre- and post-questionnaire, the reliability estimates were r = .79 and r = .92, respectively. Finally, the reliability estimate for the three items in the pre-questionnaire regarding robots broadcasting images on the Web was r = .80. As can be seen, the reliability of the service robot questions in the post-questionnaire fails to meet Nunnally and Bernstein's recommendations. The implication is that finding a difference between pre- and post-attitudes towards service robots may be threatened. However, as will be seen in the results section, despite this threat, a significant difference was found. Had the reliability of these items been larger, the difference would more likely be larger [67].

The internal consistency estimate for the personality scale was r = .74. However, when a test, such as the personality measure used in the current study, measures multiple dimensions, lower reliability estimates are expected. Furthermore, Nunnally and Bernstein (1994) assert that estimates as modest as $r = .70$ are sufficient when estimating the relationships between variables. The purpose of the personality scale was to determine the relationship between personality and attitudes towards service robots, social robots, and reactions to Cherry. Pearson-product correlation coefficients were estimated in order to determine these relationships. The major implication is that the resulting relationships may be larger if the test were more reliable. When estimating correlation coefficients, *r*- and *p*-values are estimated. *R*-values indicate the degree of relationship between variables. For more information on correlation coefficients, see [68]. *P*-values will be discussed shortly. Before the correlation coefficients were estimated, principal component analysis (PCA, a data reduction technique that finds the underlying dimensions of a test) was conducted in order to confirm that the personality items indeed did assess five aspects of personality.

The final two statistical techniques used were analysis of variance (ANOVA) and *t*-tests. These procedures allow for comparisons of mean scores between groups and/or pre- and post-events in order to determine if they are statistically different. ANOVA results in *F*- and *p*-values. *T*-tests result in *t*- and *p*-values. In both cases, the *p*-value is the probability of obtaining a particular *F*- or *t*-value if there were *no* differences between groups and/or pre- and post-events. In the behavioral sciences, in order to conclude that there is a difference between mean scores, a *p*-value equal to or less than $p = .05$ is recommended [68]. In other words, a *p*-value of $p = .05$ suggests that there is a five percent chance that the mean scores are equal, indicating that the mean scores are probably different. The same logic can be applied to correlation coefficients: a *p*-value of $p = .05$ indicates that there is a five percent chance that the resulting coefficient would be obtained if there were *no* relationship between the variables, indicating that there is probably a relationship between the two variables.

Results. Analysis of Variance (ANOVA) was conducted in order to determine the item-by-item differences between the sexes, races, and ages of the participants. Two items resulted in statistically different average scores. For example, the mean scores for the item *What level of experience do you have with robots?* differed by ethnicity $F(4, 50) = 2.818$, $p < .05$; however, overall, the participants did not have much experience with robots. Specifically, Asian participants ($M = 3.75$, $SD = 1.26$) had more experience with robots than any of the other ethnic groups (means and standard deviations ranged from 4.60-5.00 and .00-.68, respectively).

The results also indicated that the average scores for the item *Do you like Cherry's physical appearance?* differed significantly by sex $F(1, 54) = 4.617$, $p < .05$. Females ($M = 2.67$, $SD = .95$) liked Cherry's physical appearance more than males ($M = 3.36$, $SD = 1.28$). Table 4 lists the items, means, and standard deviations regarding Cherry's appearance and features. As can be seen, the subjects did not particularly like or dislike Cherry's appearance. However, the subjects did find her point-and-click map ($M = 2.23$, $SD = .97$), text box ($M = 2.05$, $SD = 1.02$), and search capabilities ($M = 2.05$, $SD = .95$) to be useful features. In addition, there was not a significant relationship between the importance of appearance when interacting with others and responses to Cherry's appearance ($r = -.13$, $p = .40$).

The mean scores of the three items measuring comfort with a robot taking pictures and broadcasting those images on the Web indicated that the participants were either unsure or uncomfortable. In particular, the subjects were slightly uncomfortable with having a robot with a camera at eye level broadcasting images on the Web ($M = 2.32$, $SD = 1.19$). In addition, they were unsure about having (a) the images viewed by the person(s) controlling the robot ($M = 2.96$, $SD = 1.28$) and (b) a robot with a camera mounted close to the floor (showing feet and furniture) broadcasting images on the Web ($M = 3.02$, $SD = 1.34$).

Table 5 presents the means and standard deviations for the five items that were in both the pre- and post-questionnaires. After exposure to Cherry, the participants' responses were significantly more positive for three items. The participants indicated that it was more acceptable for robots to display emotions ($t = 2.131$, $p < .05$) after meeting Cherry than they did before meeting her. In addition, interactive robots displaying positive emotions was more acceptable after meeting Cherry ($t = 5.753$, $p < .001$) than before meeting her. Finally, a robot displaying frustration/anger with obstacles ($t = 5.203$, $p < .001$) and people ($t = 3.274$, $p < .01$) interfering with the robot's tasks was more acceptable after meeting Cherry.

Table 4

Mean Scores and Standard Deviations for Items Regarding Cherry's Appearance and Features

Item	*M*	*SD*
Did you find Cherry's face to be pleasing?	2.91	1.06
Do you like Cherry's physical appearance?	2.84	1.07
Did you enjoy interacting with a robot that has a human face?	3.04	1.06
Do you like Cherry's overall appearance (e.g., physical and interface combined)?	2.89	1.07
Do you think the text box feature is helpful for understanding what Cherry says?	2.05	1.02
Do you like the video feature, which is the ability to see how your face is lining up with Cherry's camera?	2.77	1.25
Do you think it would be easy to use the point-and-click map to direct Cherry to someone's office?	2.23	.97
Do you like the search feature, which allows you to look up a person's name in order to find his/her office number?	2.05	.95

Five mean scores for the participants' responses were calculated from the items measuring: (1) pre-attitudes towards service robots in general (*M* = 2.83, *SD* = .94), (2) post-attitudes towards service robots in general (*M* = 2.54, *SD* = .68), (3) pre-attitudes towards robots with personality and emotion features (*M* = 3.11, *SD* = .82), (4) post-attitudes toward robots with personality and emotion features (*M* = 2.74, *SD* = 1.00), and (5) reactions to Cherry (*M* = 2.63, *SD* = .77). After they were introduced to Cherry, there was a significant change in the participants' attitudes towards robots. For example, after meeting Cherry, the participants responded more positively to the idea of service robots (t = 2.365, $p < .05$) and to robots with social abilities (t = 3.818, $p < .001$).

Finally, the factor structure of the 15 personality items was assessed with principal components analysis (PCA) using SPSS. Prior to conducting the analysis, the suitability of the data for PCA was assessed. Working in accordance to the recommendations of Tabachnick and Fidell [69], the correlation matrix was inspected and revealed that several coefficients were equal to or greater than .30. The Kaiser-Meyer-Oklin measure of sampling adequacy value was .64, exceeding the recommended value of .60 [70], [71] and the Bartlett's test of sphericity [72] was significant ($p < .001$), supporting the factorability of the items. PCA was subsequently conducted and revealed five factors with eigenvalues greater than 1, which explained 66% of the variance. In order to interpret the pattern of item loadings, Varimax rotation was performed.

Table 5

Mean Scores and Standard Deviations for Repeated Items

	Pre		Post	
Item	*M*	*SD*	*M*	*SD*
Do you think giving a robot a personality is a good feature?	2.79	1.28	2.73	1.21
Which communication method would you prefer a robot use to inform you about difficulties it is having while accomplishing tasks?	2.48	.74	2.54	.74
Do you think that having a robot display emotions could make them more accepted into everyday roles in human life?	3.05	1.28	2.76*	1.20
Do you feel that interactive robots should display emotions, positive or negative?	3.07	1.22		
Post item referring to positive emotions			2.23^{tt}	1.18
Post item referring to negative emotions			3.02	1.43
Do you feel it would be appropriate for a robot to get angry or upset with an obstacle or person that interferes with the robot's task?	4.14			
Post item referring to obstacles			3.41^{tt}	1.37
Post item referring to persons			3.68^{t}	1.32

*p < .05. ^{t}p < .01. ^{tt}p < .001.

Table 6 presents the resulting item loadings. As can be seen, with the exception of one Agreeableness item, the items corresponding to each of the personality dimensions loaded into their respective factors.

Table 6

Factor Loadings of Personality Items

Item	Factors 1	2	3	4	5
E1	.695				
E2	-.689				
E3	.813				
N1		.772			
N2		-.663			
N3		-.833			
A1			.449		
A2			.747		
A3			-		
O1				-.376	
O2				.820	
O3				.802	
C1					.800
C2					.679
C3					-.515

Note. E = Extroversion. N = Neuroticism. A = Agreeableness. O = Openness to Experience. C = Conscientiousness.

Once the factor structure of the personality items was confirmed, the three items for each personality dimension were summed and averaged. Pearson-product correlations were calculated in order to determine the relationships between the personality dimensions and five item clusters (i.e., pre- and post-attitudes towards service robots and social robots, and reactions to Cherry). One personality dimension, Openness to Experience, demonstrated a significant relationship. Specifically, Openness to Experience was negatively related to the subjects' opinions of Cherry (r = -.321, p < .05). In other words, the subjects who were more open to experience responded more positively to Cherry than individuals who were less open to experience.

Discussion. The survey revealed significant results regarding sex, ethnicity, and personality with respect to Cherry and prior experience with robots. The most significant finding with respect to sex differences was that females found Cherry's physical appearance more pleasing than males; however, there were no sex differences with regards to Cherry's avatar. It is also interesting to note that, while participants had little experience with robots, the Asian participants had more experience than any of the other ethnic categories. Because those in this study, and even more generally most people, have little experience with robots, it is important to develop robots in such a way that people will be willing to use and interact with them, or at least be open to new ideas with robotics. In fact, the results suggest that individuals who are more open to experience indeed do react more positively to robots. The results from this study also showed that exposure to Cherry changed opinions concerning social robots. As a whole, people were more open to robots displaying emotions after interacting with Cherry than before, especially with respect to robots displaying positive emotions. Although there was a more positive reaction to robots exhibiting negative emotions towards obstacles and people after exposure to Cherry, the participants still did not find it suitable.

Because of the design of Cherry, broadcasting images is essential if the operator is to be able to safely control her. Therefore, this study also aimed to determine how comfortable people would be with the use of cameras. In general, the participants were not comfortable with the use of cameras at eye-level broadcasting to the Web for many to see and not sure about how they felt about an eye-level camera viewed by only the operator or about a floor-level camera broadcasting to the Web. However, these questions were asked in the pre-questionnaire and perhaps a better time to ask them would be in the post-questionnaire, after seeing what exactly the cameras project.

As far as usability of Cherry, the participants in the study were pleased with her complete interface. The results for the survey items that referred to the text box, point and click map, and the search feature reinforced the decision to include these elements. Even though there was a negative reaction in general to the use of cameras, the participants did find the video feature used for facial recognition to be useful.

Limitations. A limitation of the survey in particular was that the reliability of the post-questionnaire items referring to service robots was low and one of the personality items referring to Agreeableness did not fall into its respective factor. In addition, because the study will be an ongoing endeavor, improvements to the scale items will be made. Therefore, more substantial positive increases in attitudes towards service and social robots as well as reactions to Cherry might be found.

VI. FUTURE RESEARCH GOALS

A. Survey Research with Cherry

As noted previously, participants from study two were predominantly from the psychology department. Further studies will incorporate people from other disciplines in order to study how background, in addition to sex and ethnicity, might influence views and reactions to Cherry. Another area of interest is the effect of age, especially with respect to individuals over 40. Previous research in the field of training indicates that older individuals may be more apprehensive towards technology than younger individuals. For example, researchers have found that older individuals report more anxiety towards technology and less confidence in their ability to learn new technology than younger individuals [73], [74], [75]. In addition, in a training program for a new technical tool, the findings suggested that older individuals found the technology to be less useful than younger individuals [76]. By expanding our pool of participants to include older individuals, we will be able to better determine whether Cherry's design and features is acceptable to a wider variety of individuals.

B. Avatar Research

Another area of concern is the importance of the use of a face, or avatar, for service and social robots with respect to interaction, usability, and understanding from a human's point of view. In the study where Bruce and colleagues [23] monitored the time students interacted with their robot, they reported that students interacted longer with the robot when it displayed a face. The authors concluded that a robot with a face is important for social robotics. However, the responses to Cherry and Lola's face in study one, described previously, indicated that the *appearance* of that face may also influence the human-robot interaction. Therefore, future work with Cherry will build on the importance of a face for human-robot interaction, the importance of physical attractiveness of the avatar, and the usefulness of an avatar for communication.

C.. More Sophisticated Personality for Cherry

Our plan is also to create a framework that enables designers to set an overall encompassing personality parameter that can predispose an agent to a specific personality type also linked with a specific set of emotions (e.g. agent with a meek personality might get discouraged more easily and give up in the face of adversity, whereas another one with an aggressive personality will get ANGRY and be inclined to fight back).

With robots collaborating with humans in a team, matching agent personality types to team members might bring about better overall group performance.

D. More Refined Emotions and Expressions of Emotions

We plan to enhance the emotion-based architecture to fully implement the AKR scheme described in [50] and to enable more sophisticated robot decision-making based on more complex emotion-like states.

In human-human communication comes from the congruency of all the various communication signals together. One can get an uncomfortable sense from an interlocutor by perceiving (consciously or not) that his or her multimodal expressions are not in sync with each other (e.g., facial expressions are incongruent with vocal intonation and body posture). In robots, similar intuitive "body" languages such as camera tilt, navigation speed, etc. can be used to exteriorize internal states to the user in a manner in which the user will naturally understand.

E. Porting The Design to a New Hardware Platform

We are currently porting the interface and the collection of social behaviors from our original toy amigobot to our new ActivMedia Peoplebot - a much more versatile robot.

F. Realistic Test beds and Applications

As mentioned before, many applications involving human-robot interaction may not benefit from including social intelligence in the robot portion of the interaction. However, some applications intuitively lend themselves to it, such as personal care (e.g., home elderly care), service robots (e.g. office assistant), and entertainment robots (e.g. toys, pets, museum docents).

Indeed, "Within a decade, robots that answer phones, open mail, deliver documents to different departments, make coffee, tidy up and run the vacuum could occupy every office" [77].

The question as to whether military robotic forces might also benefit from robots with social intelligence may not be as intuitive and might require more inquiry. These kinds of applications are very likely to depend on the type of numeric relationships and authority relationships [12].

ACKNOWLEDGEMENTS

Dr. Lisetti would like to thank *ActivMedia* for donating an Amigobot to the USF-UCF joint team winner of the AAAI 2000 *Nils Nilsson Award for Integrating AI Technology*. Thanks to Dr. Robin Murphy, Dr. Erika Rogers, and the Steering Committee for inviting Dr. Lisetti to participate to the *DARPA/NSF* Study on Human-Robot Interaction. Dr. Lisetti also would like to acknowledge the great insights received from Dr. Ian Horswill regarding the state-of-the-art in robotics. Special thanks from the authors to Rob Traub, without whose ongoing assistance, the project would have never succeeded.

REFERENCES

[1] Murphy, R. R. Introduction to AI Robotics. Cambridge, MA: MIT Press, 2000.

[2] Brooks, R and Flynn, A. "Fast, Cheap, and Out of Control", *AI Memo 1182*, MIT AI Laboratory, 1989.

[3] Arkin, R. C. *Behavior-Based Robotics*, Cambridge, MA: MIT Press, 1998.

[4] Murphy, R. R. Use of Scripts for Coordinating Perception and Action, In *Proceedings of IROS-96*, 1996a.

[5] Murphy, R. R. Biological and Cognitive Foundations of Intelligent Sensor Fusion, *IEEE Transactions on Systems, Man and Cybernetics*, 26 (1), 42-51, 1996b.

[6] Murphy, R. R. Dempster-Shafer Theory for Sensor Fusion in Autonomous Mobile Robots, *IEEE Transactions on Robotics and Automation*, 14 (2), 1998.

[7] Horswill, I. Polly: A Vision-Based Artificial Agent, In *Proceedings of the AAAI Conference*, 1993.

[8] Breazeal, C. and Velasquez, J. Toward Teaching a Robot "Infant" using Emotive Communication Acts. In *Proceedings of the Socially Situated Intelligence Workshop*, 1998.

[9] Breazeal, C. and Scasslati, B. Infant-like Social Interactions between a Robot and a Human Caretaker. *Adaptive Behavior*, 8(1), 47-72, 1998.

[10] Velasquez, J. From Affect Programs to Higher Cognitive Emotions: An Emotion-Based Control Approach. In *Proceedings of the Emotion-Based Agent Architecture Workshop at theInternational Conference on Autonomous Agents*, 1999.

[11] Murphy, R. R., Lisetti, C. L., Tardif, R., Irish, L., and Gage, A., Emotion-Based Control of Cooperating Heterogeneous Mobile Robots, *IEEE Transactions on Robotics and Automation*, Vol. 18(5), 744-757, 2002.

[12] Rogers, E. and Murphy, R. 2002. Human-Robot Interaction. *Final Report for DARPA/NSF Study on Human-Robot Interaction*, www.csc.calpoly.edu/~erogers/HRI/HRI-report-final.html.

[13] Simon, H. Motivational and Emotional Controls of Cognition. (Originally published in 1967) Reprinted in *Models of Thoughts*, Yale University Press, 1979, 29-38.

[14] Sloman, A. "Motives Mechanisms and Emotions". *Cognition and Emotion*. 1, 3, pp 217 – 234. 1987.

[15] Barbalet, J. M. *Emotion, Social Theory, and Social Structure*. Cambridge, MA: Cambridge University Press, 1998.

[16] Rosalind Picard. Potential Concerns Chapter. *Affective Computing*, Cambridge, Mass.: MIT Press, 1997.

[17] El-Nasr, M. Loerger, T. and Yen, J. "Peteei: A Pet with Evolving Emotional Intelligence. In *Proceedings of the International Conference on Autonomous Agents*, 1999.

[18] Velasquez, J. Cathexis, a Computational Model for the Generation of Emotions and their Influence in the Behavior of Autonomous Agents. Master Thesis, MIT, 1996.

[19] Michaud, F., Audet, J., Letourneau, Lussier, Theberge-Turmel, C. Caron, S. Artificial Emotions and Social Robotics. In *Proceedings Distributed Autonomous Robotic Systems*, 2000.

[20] Michaud, F., Audet, J., Letourneau, Lussier, Theberge-Turmel, C. Caron, S. Experiences with an Autonomous Robot Attending AAAI. *IEEE Intelligent Systems*, September-October, 2001.

[21] Bradshaw, J. M., "An Introduction to Software Agents", in Software Agents, Bradshaw, J.M. (ed.), Cambridge, MA: MIT Press, 1997

[22] Maybury, M. and Wahlster, W. Intelligent User Interfaces: An Introduction. *Readings in Intelligent User Interfaces*. Morgan Kaufman Publishers, 1998.

[23] Bruce, A., Nourbakhsh, I., Simmons, R. The Role of Expressiveness and Attention in Human-Robot Interaction. *Proceedings of ICRA*, 2002.

[24] Wahlster, W., Reithinger, N., and Blocher, A. SmartKom: Towards Multimodal Dialogues with Anthropomorphic Interface Agents. In: Wolf, G., Klein, G. (eds.), *Proceedings of International Status Conference on Human-Computer Interaction*, 2001.

[25] Roy, D. and Pentland, A. Adaptive Multimodal Interfaces, In *Proceedings of the Workshop on Perceptual User Interfaces*, Banff, Canada, 1997.

[26] Thalmann, N., Kalra, P., and Pandzic, I. Direct Face-to-Face Communication Between Real and Virtual Humans, 1998.

[27] Kling, R. What is Social Informatics and Why Does it Matter? *D-Lib Magazine*, 5(1), 1999.

[28] Simonson, J. and Kensing, F. Using Ethnography in Contextual Design. Communication of the ACM, Vol. 40 (7), 82-88, 1997.

[29] Clancey, W.J. (2002) Simulating activities: Relating motives, deliberation, and attentive coordination (*Cognitive Systems Research* 3(3) 471-499, September, special issue on situated and embodied cognition).

[30] Dautenhahn, K. Robots as Social Actors: Aurora and the Case of Autism. *Proceedings of the Third Cognitive Technology Conference*, 1999.

[31] Eysenck, H.J. (1991). Dimension of personality: 16, 5, or 3? – Criteria for a taxonomic paradigm. *Personality and Individual Differences, 12,* 773-790.

[32] Costa, P.T., Jr., & McCrae, R.R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653-665.

[33] Cattell, R.B., Eber, H.W., & Tatsuoka, M.M. (1970).

Handbook for the sixteen personality factor questionnaire. Champaign, IL: Institute for Personality and Ability Testing.
[34] Gough, H.G. (1987). *California Personality Inventory Administrator's Guide.* Palo Alto, CA: Consulting Psychologists.
[35] Saville, P., Holdsworth, R., Nyfield, G., Cramp, L., & Mabey, W. (1984). *The Occupational Personality Questionnaires (OPQ).* London: Saville and Holdsworth (UK) Ltd.
[36] Matthews, G., & Deary, I.J. (1998). The trait concept and personality theory. In *Personality Traits* (pp. 3-37). New York, NY: Cambridge University Press.
[37] Digman, J.M., & Takemoto-Chock, N.K. (1981). Factors in the natural language of personality: Re-anlaysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research, 16,* 149-170.
[38] Fiske, D.W. (1949). Consistency of the factorial structures of personality ratings from different sources. *Journal of Abnormal and Social Psychology, 44,* 329-344.
[39] Goldberg, L.R. (1990). An alternative 'Description of personality': The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*, 1216-1229.
[40] Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66,* 574-588.
[41] Ostendorf, F., & Angleitner, A. (1994). Reflections on different labels for Factor V. European *Journal of Personality, 8,* 341-350.
[42] Tupes, E.C., & Christal, R.E. (1961). Recurrent personality factors based on trait ratings. Technical Report No. ASD-TR-61-97, US Air Force, Lackland US Air Force Base. *TX/Journal of Personality, 60,* 225-251.
[43] Webbs, E. (1915). Character and intelligence. *British Journal of Psychology Monographs, I, 3,* 1-99.
[44] Krug, S.E., & Johns, E.F. (1986). A large scale cross-validation of second-order personality structure defined by the 16PF. *Psychological Reports, 59,* 683-693.
[45] H. Leventhal and K. Scherer, "The relationship of emotion to cognition: A functional approach to a semantic controversy," *Cognition and Emotion,* vol. 1, no. 1, pp. 3 – 28, 1987
[46] Zajonc R. Feeling and Thinking: Preferences Need No Inferences. American Psychologist, 35(2), 1980.
[47] Frijda, N.H., *The Emotions.* New York: Cambridge Univ. Press, 1986, vol 8.
[48] Ortony, A. On Making Believable Emotional Agents Believable . In *Emotions in humans and artifacts*, R. Trappl & P. Petta P. (eds.), MIT Press, Cambridge, 2001.
[49] Wierzbicka, A. *Emotions Across Languages and Cultures*, Cambridge University Press. 1999.
[50] Lisetti, C. L. Personality, Affect and Emotion Taxonomy for Socially Intelligent Agents. In *Proceedings of the Florida International Artificial Intelligence Society Conference.* Menlo Park, CA: AAAI Press, 2002a.
[51] Lisetti, C.L and Bianchi, N. Modeling Multimodal Expression of User's Affective Subjective Experience. User Modeling and User-Adapted Interaction, An International Journal, *Vol. 12(1): 49- 84. 2002.*
[52] Hayes-Roth, B., Ball, G., Lisetti, C. L. Picard, R., Stern, A. Panel on Affect and Emotion in the User Interface. In *Proceedings of the International Conference on Intelligent User Interfaces*. New-York, NY: ACM Press, 1998.
[53] ActivMedia. www.activmedia.com, 2002.
[54] Identix, www.identix.com, 2002.
[55] Haptek, www.haptek.com, 2002.
[56] Lisetti, C. L. and Schiano, D. J. Automatic Facial Expression Interpretation: Where Human-Computer Interaction, Artificial Intelligence, and Cognitive Science Intersect. *Pragmatics and Cognition*, 8(1), 2000.
[57] Darwin, C. "The expression of the emotions in man and animals". vi,372. D. Appleton and Company. New York. 1898
[58] Ekman, P., & Rosenberg, E.L. (editors). What The Face Reveals: Basic And Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). New York: Oxford University Press, 1997
[59] Rinn, W.E. "The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions". *Psychological Bulletin*. 95. 52-77. 1984.
[60] Chovil, N. Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction* 25:163–194. 1991.
[61] Fridlund, A.J. "Human facial expression: An evolutionary view". NY. Academic. 1994.
[62] Camras, L. "Expressive development and basic emotions". *Cognition and Emotion.* 6. 269 – 283. 1992.
[63] Ekman, P. and Davidson, R. The Nature Of Emotion: Fundamental Questions. New York: Oxford University Press, 1994.
[64] Zajonc, R., and Markus, H. Affect and Cognition: The Hard Interface. In Izard et al. (Eds.) *Emotions, Cognition and Behavior*, 73-102, 1984.
[65] Lisetti, C. Motives for Intelligent Agents: Computational Scripts for Emotion Concepts. In Proceedings of the Sixth Scandinavian Conference on Artificial Intelligence (SCAI'97), Amsterdam, Holland: IOS Press Frontiers in Artificial Intelligence and Applications, (Helsinki, Finland, August 1997).
[66] Lisetti, C.L. and Gmytrasiewicz, P. Can a Rational Agent Afford to be Affectless? A Formal Approach, Applied Artificial Intelligence, Vol. 16, 577-609, 2002.
[67] Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric Theory* (3rd ed.). New York, NY: McGraw-Hill.
[68] Shavelson, R.J. (1996). *Statistical Reasoning for the Behavioral Sciences* (3rd ed.). Needham Heights, MA: Allyn and Bacon.
[69] Tabachnick, B.G., & Fidell, L.S. (2001). *Using Multivariate Statistics* (4th ed.). Needham Heights, MA: Allyn and Bacon.
[70] Kaiser, H. (1970). A second generation Little Jiffy. Psychometrika, 35, 401-415.
[71] Kaiser, H. (1974). An index of factorial simplicity. Psychometrika, 39, 31-36.
[72] Bartlett, M.S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16* (Series B), 296-298.
[73] Colquitt, J.A., LePine, J.A., & Noe, R.A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*, 678-707.
[74] Martocchio, J.J. (1994). Effects of conceptions of ability on anxiety, self-efficacy, and learning in training. *Journal of Applied Psychology, 79*, 819-825.
[75] Salanova, M., Grau, R.M., Cifre, E., & Llorens, S. (2000). Computer training, frequency of usage and burnout: The moderating role of computer self-efficacy. *Computers in Human Behavior, 16*, 575-590.
[76] Warr, P., Allan, C., & Birdi, K. (1999). Predicting three levels of training outcome. *Journal of Occupational & Organizational Psychology, 72*, 351-375.
[77] Lorek, L, "March of the A.I. Robots", *Interactive Week,* April 30. 2001

Christine Laetitia Lisetti received the B.S degree in Computer Science c*um laude*, and the M.S. and Ph.D. degrees in computer science in 1988, 1992, and 1995, respectively, from Florida International University, Miami. From 1996 to 1998 she was a Post-doctoral Fellow at Stanford University, jointly in the Department of Computer Science and the Department of Psychology, Stanford.

She is currently an Assistant Professor in Computer Science at the University of Central Florida, Orlando. Her research focus on affective computing and include affective-cognitive process modeling, intelligent affective user interfaces and human-robot interaction. Dr. Lisetti has won multiple awards including a National Institute of Health Individual Research Service Award, and the AAAI Nils Nilsson Award for Integrating AI Technologies.

Sarah M. Brown received her B.S. in computer science from the University of Central Florida in 2002. She is currently pursuing her M.Sc. in computer science from Simon Fraser University in Burnaby, BC. Her research interests include affective computing, intelligent agents, and user interfaces.

Kaye Alvarez received the M.S degree in Industrial-Organizational Psychology from the University of Central Florida in August, 2003. Currently, she is a Test & Measurement Technician for the Personnel Board of Jefferson County, Birmingham, Alabama. Her research interests include test development and psychometrics, emotional intelligence, affective computing, and personnel selection and training.

Andreas H. Marpaung received the B.S degree in Computer Science c*um laude* from the University of Central Florida in 2002, where he is currently pursuing a Computer Science Ph.D. degree. His research interests include affective computing, artificial intelligence, and intelligent agents.

APPENDIX B: Publication II

S. Brown, C. Lisetti and A. Marpaung. Cherry, the Little Red Robot with a Mission and a Personality. In *Working Notes of the AAAI Fall Symposium Series on Human-Robot Interaction*, Menlo Park, CA: AAAI Press. Cape Cod, MA. November 2002.

Cherry, The Little Red Robot…
with a Mission … AND a Personality!

Sarah M. Brown, Christine L. Lisetti, Andreas H. Marpaung

Computer Science Department
University of Central Florida
Orlando, Florida 32816
sbrown@cs.ucf.edu; Lisetti@cs.ucf.edu; marpaung@cs.ucf.edu

Abstract

In this article, we describe the development of an autonomous robot, Cherry the Little Red Robot, whose functionality we designed so that she could socially interact with humans on a daily basis in the context of an office suite environment. Cherry has a given set of office-tasks to accomplish, from giving tours of our computer science faculty suite to visitors, to serving beverages to those faculty and staff, and to engaging them in social interaction. We describe (1) our motivation for social informatics in human-robot interaction, (2) the hardware additions that we implemented for ActivMedia robot for our purposes, as well as (3) the Cherry's multi-modal anthropomorphic interface that we developed capable of combining speech, face recognition, and emotional facial displays, and finally (4) our future research efforts.

Introduction

With increasing advances in robotics in behavior-based robotics (Brooks, 1989; Arkin, 1998), sensor fusion (Murphy, 1996a, 1996b, 1998, 2000), robot vision (Horswill, 1993), emotion-based architectures (Breazeal, 1998; Velasquez, 1999; Murphy, Lisetti et al., 2002), a variety of domains and applications for human-robot interaction and collaboration are emerging: planetary exploration, urban search and rescue (USAR), military robotic forces, personal care and service robots (e.g. hospital assistance, home elderly care, robotic surgery), home appliances, entertainment robots, and more (Rogers and Murphy, 2002).

Furthermore, recent progress in artificial intelligence, speech simulation and understanding, graphics and computer vision have made it possible to design computer systems that have "social expertise" in order to naturally bring the human – a principally social animal (albeit engineering formal training has altered natural preferences) – into the loop of human-computer interaction. Social informatics has indeed been considered a critical, unexplored area in the domain of human-robot interaction (Rogers, Murphy, 2002) which we currently set out to explore and to focus our contribution on.

In this article, we consider social expertise in terms of (1) external communicative behavior and (2) internal motivational goal-based abilities (Lisetti, 2002). We describe the development of a service application on an autonomous robot, Cherry the Little Red Robot shown in Figure 1. Cherry's functionality has been designed so that she could socially interact with humans on a daily basis in the context of an office suite environment.

Figure 1: Cherry, our AAAI 2000 Prize

Cherry has a given set of office tasks to accomplish, from giving tours of our computer science faculty suite to visitors, to serving beverages to those faculty and staff and to engaging them in social interaction. In the remainder of this article, we describe:

(1) our motivation for focusing our design, hardware, and user interface on social informatics and emotional intelligence,
(2) the basic hardware included in our ActivMedia robot as well as the hardware additions that we implemented for our purposes,
(3) Cherry's multi-modal anthropomorphic interface capable of combining speech, face recognition, and emotional facial displays so that she is socially engaging to humans from the very start of her functional design and implementation,

(4) and finally our future software developments to complete our project.

Relevant Background on Human-Robot Interaction and Social Informatics

Ten years ago, it was predicted that robots would become important factors in home and office environments (Ralston, 1993). As documented in the Final Report for DARPA/NSF Study on Human-Robot Interaction (Rogers and Murphy, 2002), although complete robot autonomy has not yet been accomplished, "the feasibility of integrating various robot entities into people's daily lives is coming much closer to reality. […] robots now have the potential to serve not only as high-tech workhorses in scientific endeavors, but also as more personalized appliances and assistants for ordinary people."

However, it has also been noted (Rogers and Murphy, 2002) that before autonomous and intelligent robots are fully integrated into our society, the nature of human-robot relationships and the impact that these relationships may have on our future need to be very carefully considered.

Indeed, robots differ from simple machines or computers in that they are mobile, have varying levels of autonomy, and therefore are not as predictable, and can furthermore interact within a user's personal physical space. When such a robot has autonomy, the social interaction that results is unlike any previous man-machine relationships.

From our perspective, an interesting modeling issue therefore becomes that of social relations. In particular, we have chosen to focus our contribution to the field in addressing the technical goals of (1) understanding how to embody affective social intelligence and of (2) determining when embodied affective social intelligence is useful. Toward that goal, we have identified a collection of relevant questions and we have categorized them into three main categories discussed below:

1. Robot social intelligence: for example, can "no personality" in an intelligent agent (software or robot) be perceived by humans as a cold, insensitive, indifferent agent? If so, do these perceptions differ by specific groups of people, differentiated by age, gender, culture, etc.? Is it important to change the perceptions mentioned above in humans so that agents can be viewed as personable, helpful, maybe even compassionate? If such is the case, can we identify the various contextual situations and applications when that is beneficial, or even necessary? If emotions and personality are embodied in a robot, does it affect how the people respond to it? If so, how so, and in what contexts? Should they resemble that of humans, or should they be depart from them?

2. Human social intelligence: on the other hand, one may also ask how do the personality of the human affect how the human interacts with the robot? If so, how? Does it arouse specific emotions, behaviors? Which ones? In what contexts does this happen? Are these effects consistently observable, predictable, positive, or negative? Can we improve on these toward the positive? How so?

3. Human-Robot social relationship: finally, looking at the relationships themselves, questions arise as to what kind of taxonomy of human-robot social "relationships" can be established, identifying numeric (eg. 1:1, 1:m, m:m), special (eg. remote, robo-immersion, inside), and authority (eg. supervisor, peer, bystander) relationships (Rogers and Murphy, 2002) to determine what levels of "interpersonal skills" a robot would need in order to perform its role effectively.

Affective Social Intelligence

In order to understand when these social relationships are needed or when the perception of such relationships need to be changed, social relations must be modeled. Emotions have an evolutionary crucial functional aspect in intelligence without which complex intelligent systems with limited resources cannot function nor behave efficiently (Simon, 1967).

Emotions are carriers of important messages which enable an organism to maintain a satisfactory relationship with its environment. *Fear*, for example, serves the function of preparing an organism physiologically for a flight-or-fight response (blood flow increases to the limbs, attentional cues are restricted, etc.). *Anxiety*, on the other hand, serves the function of indicating that further preparation for the task at hand is needed. Other examples of the functions of emotions abound (Lisetti, 2002).

Emotions greatly influence decision making (although sometimes dysfunctionally), more often than not for improved efficiency and flexibility toward a complex changing environment (Lisetti and Gmytrasiewicz, 2002). Indeed, pure reasoning and logic have proven to be insufficient to account for true intelligence in real life situations. In the real world with all its unpredictable events for example, there is not always time to determine which action is best to choose, given an infinite number of possible ones and a set of premises.

Furthermore, different personalities will incline individuals to have different mental and emotional pattern tendencies. An agent with an *aggressive* personality, for example, will be predisposed to a *fight* response when experiencing fear, whereas one with a *meek* personality will be predisposed to *flee*. Predispositions, however, can be altered by conscious repression and/or adaptation.

Furthermore, **personality predisposes** an agent toward a certain set of emotional states and action tendencies: We consider personality as representing characteristics of an autonomous self-motivated organism that account for consistently chosen patterns of reaction over situations and time including behavior, emotions, and thoughts.

Relevant Applications for Social Human-Robot Interaction

As mentioned before, many applications involving human-robot interaction may not benefit from including social intelligence in the robot portion of the interaction.

However, some applications intuitively lend themselves to it, such as personal care and service robots (e.g. home elderly care, office assistant), entertainment robots (e.g. toys, pets, museum docents). Indeed, "Within a decade, robots that answer phones, open mail, deliver documents to different departments, make coffee, tidy up and run the vacuum could occupy every office".

Military Applications

The question as to whether military robotic forces might also benefit from robots with social intelligence may not be as intuitive and might require more inquiry. These kinds of applications are very likely to depend on the type of numeric relationships and authority relationships (Rogers and Murphy, 2002).

For certain types of applications, modeling emotions and personality of robots, agents, and humans is therefore crucial to:

- render the robots/agents more efficient themselves in terms of self-motivation, monitoring progress toward their goals, and adapt their behavior flexibly to unpredictable environments;
- work with and train humans in a more realistic environment in team work where robots can embody personality traits and emotion-like states to provide test-bed for adaptation/learning to specific personality types, emotional coping behaviors.
- predict team behaviors in terms of likelihood of task success/failure given specific mixes of agent personality types (e.g. team consisting of aggressive members only vs. team consisting of ½ aggressive and ½ meek members, altruistic vs. selfish), external environmental inputs (e.g. high stress vs. low stress, various drugs), internal individual beliefs (e.g. self-confidence levels), various emotions and moods (e.g. discouragement vs. anger).

The Office Assistant Application: Cherry's Job

In order to begin our inquiry on the modeling aspect of human-robot social relationships, we identified one specific application which we believe is intuitively enough "social" to start generating interesting relevant results.

Cherry, our little red robot, is being designed and programmed to have a variety of social roles to include being a gopher for the department and giving tours of the building.

In addition, Cherry is also being designed to have a variety of internal states and external behaviors such as:

(1) maintaining and expressing a consistent personality throughout the series of interactions;
(2) experiencing different inner emotional-like states in terms of her progress toward her goals;
(3) choosing (or not) to express these inner states in an anthropomorphic manner so that humans can intuitively understand them;
(4) having an internal representation of her social status as well as the social status of her "bosses".
(5) adapting to the social status of the person she is interacting with by following acceptable social etiquette rules.

Furthermore, to evaluate Cherry's performance and perception by humans, both during and after implementing Cherry's mission and personality, we are conducting surveys, questionnaires and experiments to begin to answer the three categories of questions mentioned earlier (human social intelligence, robot social intelligence, human-robot social interaction).

Introducing Cherry, the Little Red Robot

We won Cherry (anthropomorphically named for her social role), shown in Figure 1, at the AAAI Mobile Robot Competition entitled *Hors D'Oeuvres Anyone?* where our joint USF-UCF entry consisted of two heterogeneous human-sized cooperating Nomad robots serving hors d'oeuvres at the conference main reception (Murphy et al., 2002). Contrary to the Nomads, Cherry is a very small robot of the ActivMedia AmigoBot family (ActivMedia, 2002).

The Robot Hardware Itself

An AmigoBot is an intelligent mobile robot, capable of autonomous or user defined movement and behavior. It not only has an operating system on the robot (the AmigoOS), but is packaged with several programs that allow the user to manipulate the robot. The AmigoBot is intended for use in areas such as schools, labs, offices, and any place that is wheelchair accessible.

In our case, Cherry is intended to navigate our Computer Science Faculty Offices Suite located on the second floor of our UCF Computer Science Building. One of the main advantages of the AmigoBot is that it is highly maneuverable with 2 rubber tires, each driven by a reversible DC motor, and a rear caster for support. Furthermore, she has a red polycarbonate body that resists damage from running into obstacles.

Cherry has 8 sonars with 6 in the front and 2 in the rear (the round circles seen in Figure 1 above). Not only are AmigoBots robots able to detect if there are objects in front of, to the side of, or behind them, but to also determine how far away they are.

Human-Robot Communication and Control

The AmigoBot line also offers users the choice of connection types. Since winning Cherry, we have acquired a second AmigoBot, Lola (shown in Figure 2), which has in addition to Cherry's hardware, a camera and image transmitting device.

Figure 2: Lola featuring the antenna of the wireless modem, camera, and image transmitting device.

In addition, both of our robots have wireless capabilities provided by a pair of wireless modems for each. One modem is connected to the robot and stored underneath between the wheels, the other is connected to the serial port of a PC.

These modems have a range of approximately 300ft, but it is considerably less in areas with many walls or sharp turns. The other option for communication with the robot is via a direct connection from the PC serial port, to one inside (or on top of) the robot shown in Figure 3.

While using a tether does offer a much faster and more reliable method of transferring information, the length of the cable significantly reduces the overall uses that the robot can perform.

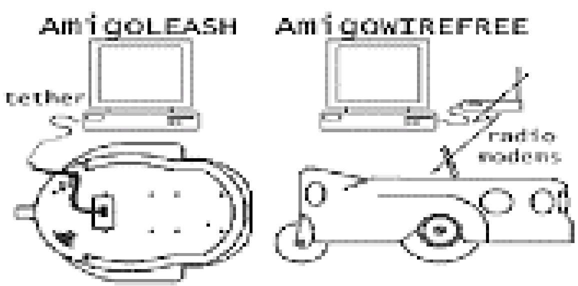


Figure 3: AmigoLeash and AmigoWireFree

Using the *ActivMedia Saphira* software (ActivMedia, 2002), we have been able to create short navigational programs for Cherry, demonstrating her motor precision and turning ability. While these programs have been useful for demonstration purposes and to work with the way in which the robot rotates, the overall usefulness for navigating the Computer Science Building is minimal.

More "Brain Power" for Cherry: Hardware Additions

Indeed straight from the factory, the Amigobot sensors are powered and processed from a single controller, driven by a high-performance, I/O rich 20MHz Hitachi H8 microprocessor. Even though acceptable for short navigations, in order to build a meaningful system with multimedia abilities, we equipped Cherry with additional processing power.

Because the robot is able to carry a payload of up to 2 lbs, we mounted a small and light laptop on top of it, which not only provides the direct connection to the robot but also the mobility since it too runs off of battery power. We equipped Cherry with a Sony Vaio Picturebook laptop in order to boost the processing power of the robot (see Figure 4).

Because the laptop is directly connected to the robot, there is no loss of data with commands as there can be using wireless modems.

However, because the laptop is designed to be extremely small, many accessories ports, such as serial and network, are either nonexistent or only present when using a port replicator. Because the robot requires a serial connection, we had to "create" a serial port using a USB to serial converter, and then use a serial cable to connect to the robot.

Another significant hardware obstacle was the addition of an OrangeMicro iBot, a FireWire camera that required a 6-pin connection with the laptop. Since our compact laptop only had a 4-pin connection, we used a FireWire hub and modified it to draw power off of the robot's battery to link the camera to the computer.

All of these cables made the robot look unsightly so we had a mount created for Cherry out of a honeycomb aluminum product that would be strong without adding much weight. This way, all of the cables are now tucked underneath the mount and out of sight: looks are important for social interactions… Furthermore the mount also provides a platform for the laptop to be mounted onto, as well as a base for installing a camera.

Indeed, because Cherry is interesting to us principally for social interaction in an office suite, we wanted to be able to mount a camera at human-eye level on the robot, to enable it to process people's faces with computer vision algorithms (described later). The final result of our hardware configuration is shown in Figures 4 and 5.

Figure 4: Cherry with laptop

Figure 5: Cherry with platform and eye-level camera

Not only has the laptop boosted the capabilities of the Amigobot, but it has also opened new doors for controlling Cherry's whereabouts.

Cherry the Little Red Robot … With a Mission…

As mentioned earlier, our current interest working with Cherry is to involve her in social interactions in an office environment while performing useful simple tasks such as, in our current project, giving a tour to computer science visitors by introducing professors and research interests as she accompanies visitors down the faculty suite, or delivering beverages and documents to professors and staff, etc.

Cherry, our little red robot, is being designed and programmed to have a variety of social roles:

(1) her master's favorite office gopher (a 1-1 master-slave human-robot relationship);
(2) her department favorite gopher (a many-1 masters-slave human-robot relationship);
(3) her department tour guide (another many-1 human-robot relationship).

We used *ActivMedia Mapper* software to create a map of our Computer Science office suite in order to (1) have the ability for simple point-and-click navigation, and to (2) have a built-in grid system that we are going to be using in the navigational portion of our interface.

Computer Science Faculty Suite Map Creation

Only walls and large permanent obstacles had to be drawn in, since the robot is able to use its sonar to navigate around smaller or moving items (and soon its vision system for collision avoidance describe later under future development).

To draw the walls, either the robot can be used to take measurements, or standard methods such as a measuring tape or blue prints can be used. We used standard methods to generate our map of the Computer Science faculty suite shown in Figure 6.

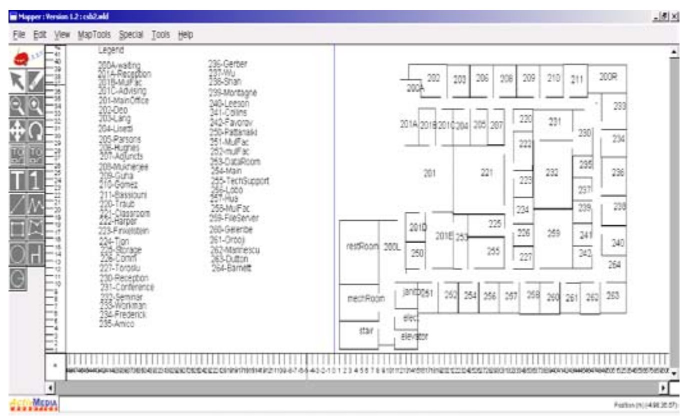


Figure 6: Map of UCF Computer Science Building Second Floor

The created map shown Figure 6, displays on the right handside each office number positioned in respectively, whereas the left handside displays the name of each professor or staff working in that office so that the user can point and click on the office to send Cherry to.

The map therefore provides quick and simple direction for Cherry. Because our map is very accurate, it also provides the basis for our (x,y) coordinate system, has been used to generate the button based map in the C++ user interface as seen in Figure 7 below, and it can furthermore be loaded in the simulator as described next.

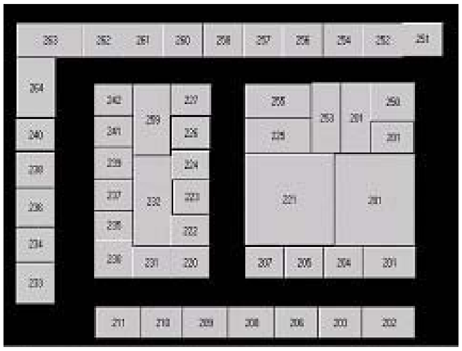


Figure 7: Button Based Map in Interface

Cherry's Navigation Simulator

The created map can also be loaded into the *ActivMedia Navigator* software (whose user interface is shown in Figure 8) in order to run and test our code in simulation rather than with the actual robot.

Indeed, given that our Affective Social Computing Lab is located on the first floor of the Computer Science building, we are making use of simulation to avoid carrying Cherry upstairs on the second floor at each step of implementation.

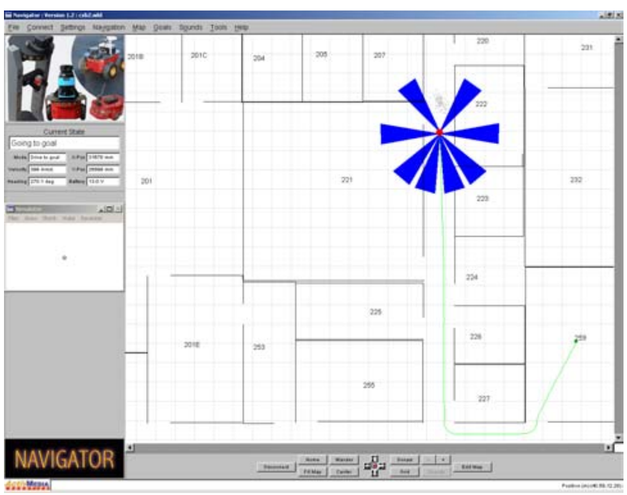


Figure 8: Navigation Simulator Interface

The *ActivMedia Navigator* software is a basic navigation module that lets you control the robot's actions with point-and-click or with menus.

As shown in Figure 8, the robot center is displayed with a red circle, the robot's sonar readings are displayed in blue and its speed and status readings (displayed on the left handside) change as it moves. The user can point on the map to the desired location, and a green dot marks it. A green line from the robot center to its final destination shows the robot path.

Web-based Interface for Cherry Command-and-Control

In order to allow users to control Cherry from their desktop (rather than having to crawl around on the floor to manipulate the laptop), we connected the laptop to the UCF network via a wireless Ethernet card.

Using WinVNC, a free piece of remote desktop control software, a user can view what is being displayed on Cherry's laptop screen and control her actions by clicking on the map interface from their desktop computer. It is important to note that WinVNC runs on all win32-based systems, and that it only requires a VNC client or a java-enabled web browser in order to view the laptop's screen. It allows both single and multiple users to connect to the laptop at our discretion. Additional users can interfere with the primary user of the laptop since they share a single view of the machine. This problem only grows larger as more people attempt to use VNC to control the robot, which is the reason we are currently restricting access only to the primary user of the robot.

We are researching better ways of displaying the web interface in order to (1) reduce potential interferences, and to (2) get a better refresh rate and color display than WinVNC can provide due to the subtle coloration and frequent movement of the avatar.

Web-based Cherry's-eye-view of the world

In addition, using the iBot camera and the wireless Ethernet capabilities of the laptop, we are able to stream video of the people and environment that Cherry encounters. This and another camera mounted nearer to her base can provide a "Cherry's-eye-view" of the world for users to access via the web.

Because an ongoing broadcast on the web of Cherry's view may not be of much interest to a large group of people, and because it would most likely raise societal privacy issues related to "Big Brother Technology", we are really using this feature solely as a very convenient way to give remote demonstrations. Future uses of the feature might emerge in the future, and will need to be cleared through subject consent forms with computer science faculty suite inhabitants.

Cherry, the Little Red Robot with a Mission… AND her Personality!!

The laptop addition also allowed us to install several other pieces of software to enhance the functionality and social behavior of our robots.

Cherry as our favorite Computer Science Tour Guide

In order to give Cherry knowledge of who works where so that she could give meaningful and instructive tours of our faculty offices, we also linked each office on the map we created with each professor or staff's facial image and current research interests (available from our UCF Computer Science web site). Now Cherry has information to introduce each person once she reaches their office.

Cherry as my favorite office gopher

We also thought that one important task for Cherry to be able to perform on our floor, was to bring soda cans to a specific professor or staff member.

First, we created a copy of the Computer Science map onto Cherry's laptop interface to enable users (for now only one) to point and click which location on the map they want Cherry to go to (see Figure 7). Users can therefore point and click on one of the map of offices drawn, which has the effect of dispatching Cherry to that office.

Figure 9: Introducing Cherry's Face

In addition, in order to facilitate the social interaction with humans, an anthropomorphic avatar has been created for Cherry to represent "her". The avatar (shown in Figure 9) is present on the laptop/Cherry's user interface and has voice ability so that she can speak to the user in natural language. She explains a variety of facts, from who she is and what her mission is, namely the UCF computer science tour guide, to which professor works in what office, to what that particular professor is researching.

Taking a social informatics co-evolutionary approach to the study and design of technology and social structures, we adopted both a bottom-up and a top-down approach to designing Cherry. We have given her social expertise in terms associating a variety of *external expressive behaviors* with her various inner states. These inner states – measured in terms of her current relationship with her environment and goals – will need to be integrated with the external behavior for a consistent system (Ortony, 2001).

From a co-evolutionary perspective, however, our bi-directional approach enables us to start testing and evaluating our interface design with human subjects while Cherry's functionality is being designed to ensure maximum success in both her functionality and interface.

Currently, Cherry can display different facial expressions corresponding to her different inner states:

- **Frustration:** Cherry expresses frustration (see Figure 10) when she finds that the office to which she was sent to has its door closed, or the door is open but she cannot recognize the faculty inside the office.

Figure 10: Cherry Frustrated

- **Discouraged**: Cherry shows discouragement (see Figure 11) when, after waiting for a while (a parameter of her patience which can be adjusted), the door remains closed.

Figure 11: Cherry Discouraged

- **Angry:** Cherry can also express anger when, after waiting for a long time, the door still remains closed (Figure 12). This option was created in order to test how people might react to her anger differently: some might want to help her accomplish her goal, whereas others might not want to deal with her at all. We plan to study these issues through psychological experiments and surveys.

Figure 12: Cherry Angry

Good Morning, Doc!

Finally, we have also integrated our face recognition system using our existing MOUE system (Lisetti, 2002) and Visionics face recognition code. Cherry now has the ability to take pictures of people she encounters, match them to our existing database of people we know have offices in the computer science suite, and greet them by name, and social status.

Part of this system enables Cherry to greet different people according to their university status. For instance, a Full Professor is greeted with more deference than a 'mere" Graduate Student, following some social rules given to her.

Complete Interface

All of the components have been integrated into a single interface designed for ease of use for both the controller of the robot and those she encounters and to provide fun, social interactions with Cherry. All of the different aspects can be seen below in Figure 13.

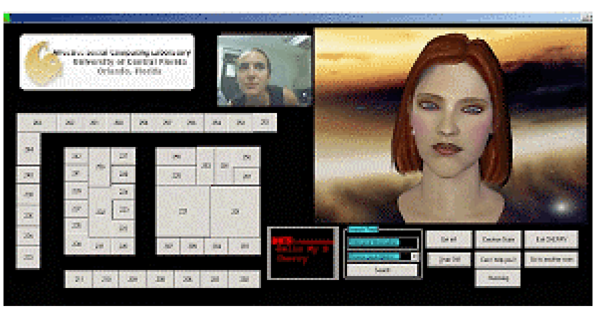


Figure 13: Cherry's User Interface

In addition to the point and click map, the avatar that is Cherry's "face", and the video components (both facial recognition and video streaming), other features have been added to assist in using the interface to control the robot. A search feature allows people to find a professor's office number and a text box displays everything that Cherry says to 1) allow hearing impaired individuals to read what is being said and 2) since remote users cannot hear anything from the laptop, they can instead read her speech. We have also added a series of buttons to allow users to ask for instructions on use, change Cherry's goals, and to exit the program.

Our Future Goals:

More Hardware and Software for Cherry: her new eyes… and vision system

To provide more autonomy for Cherry, we are going to use the Aria classes provided by ActivMedia to incorporate navigation with collision avoidance code (Horswill, 2001). Not only will we be able to dictate where she goes, but we can also use her sonar and her "eyes" to detect and avoid objects using a routine.

We also hope to orient her on a grid so that she will always "know" where she is using coordinate system. This will also assist us when attempting to have two or more robots cooperating together.

More Personality for Cherry

We also plan to create more "personality" for Cherry. Examples include: humming as she travels to and from offices, getting upset when she can't find someone in their office, getting frustrated when encountering obstacles, and getting excited when she finds who she is looking for.

Her comments and simple conversation could also in the future be tailored to individuals and their interests. We want to make Cherry interesting, lovable, and able to interact on a social level with the people who work in the building.

More Missions for Cherry

While it will be relatively simple to travel to an office with a Coke, and speak an offering, we would like to be able to have the professor select which beverage they would like. Cherry could then come back to us, tell us what she needs, and she would deliver what the professor their preferred drink.

After the "beverage offering" behavior is created and working, we plan on creating more options for the user to choose for the robot. Some of the options we have planned are: delivering a message to multiple staff members and being able to give directions to staff offices from anywhere on the second floor of the Computer Science Building.

ARIA

All of these additions will need to be developed with the ARIA software. ActivMedia Robotics Interface for Application (ARIA) is an object-oriented, robot control applications-programming interface for ActivMedia

Robotics' mobile robots. ARIA is written in C++ and provides access to the robot server as well as to the robot's sensors and accessories. It makes an excellent foundation for higher-level robotics applications, to include Saphira. ARIA is released under the GNU public license, which means that if any work you do is distributed that uses ARIA, all of the source code must be distributed. The classes that comprise ARIA are available and can be used in other C++ code, provided that it is compiled under Microsoft Visual C++.

Figure 14: Cherry, Lola, and friends

Human-Robot Social Interaction: Experiments

Finally, we plan to conduct a series of psychologically sound experiments on how humans interact, appreciate, dislike, or like our social robots.

Interestingly, for example, we have already informally established that Cherry is a very attractive for children and young college students who find her looks and autonomy absolutely irresistible.

Gender, age, personality, context and other factors are likely to influence the nature of the interaction with our social robots. We are interested in studying these issues in further details.

Acknowledgements

Dr. Lisetti would like to thank Robin Murphy, Erika Rogers and the DARPA/NSF Study on Human-Robot Interaction Steering Committee for inviting her to participate. The authors would also like to thank the *Office of Naval Research* for partial funding for this research project.

References

ActivMedia. www.activmedia.com, 2002

Arkin, R. C. *Behavior-Based Robotics*, Cambridge, MA: MIT Press, 1998.

Breazeal, C. and Scasslati, B. Infant-like Social Interactions between a Robot and a Human Caretaker. *Adaptive Behavior*, 1998.

Brooks, R and Flynn, A. "Fast, Cheap, and Out of Control", AI Memo 1182, MIT AI Laboratory, 1989.

Dyer, C. Getting Personal with Computers: How to Design Personalities for Agents. *Applied Artificial Intelligence*, 13:273-295, 1999.

Eysenck, H. Biological Dimensions of Personality. *Handbook of Personality: Theory and Research*, NY, NY: Guildord Press, 244-276.

Horswill, I. Polly: A Vision-Based Artificial Agent, In *Proceedings of the AAAI Conference*, 1993.

Lisetti, C. L. Personality, Affect and Emotion Taxonomy for Socially Intelligent Agents, In *Proceedings of the Florida International Artificial Intelligence Society Conference*, Menlo Park, CA: AAAI Press, 2002.

Lisetti, C. L. and Gmytrasiewicz, P. Can a Rational Agent Afford to be Affectless? A Formal Approach, *Applied Artificial Intelligence*, Vol. 16:577-609, 2002.

Murphy, R. R. *Introduction to AI Robotics*. Cambridge, MA: MIT Press, 2000.

Murphy, R. R. Use of Scripts for Coordinating Perception and Action, In *Proceedings of IROS-96* , 1996a.

Murphy, R. R. Biological and Cognitive Foundations of Intelligent Sensor Fusion, *IEEE Transactions on Systems, Man and Cybernetics*, 26 (1), 42-51, 1996b.

Murphy, R. R. Dempster-Shafer Theory for Sensor Fusion in Autonomous Mobile Robots, *IEEE Transactions on Robotics and Automation*, 14 (2), 1998.

Murphy, R. R., Lisetti, C. L., Irish, L., Tardif, R. and Gage, A., Emotion-Based Control of Cooperating Heterogeneous Mobile Robots, *IEEE Transactions on Robotics and Automation*, Vol. 18, 2002.

Ortony, A. On Making Believable Emotional Agents Believable . In *Emotions in humans and artifacts*, R. Trappl & P. Petta P. (eds.), MIT Press, Cambridge, 2001.

Ralston, A. and Reilly, E. D. (Eds) *Encyclopedia of Computer Science, 3rd*. Ed., Van Norstrand Reinhold: New York, 1993.

Rogers, E. and Murphy, R. Human-Robot Interaction. *Final Report for DARPA/NSF Study on Human-Robot*

Interaction, www.csc.calpoly.edu/~erogers/HRI/HRI-report-final.html, 2002

Simon, H. Motivational and Emotional Controls of Cognition. (Originally published in 1967) Reprinted in *Models of Thoughts*, Yale University Press, 1979, 29-38.

Simon, H. Making Management Decisions: The Role of Intuition and Emotion. In W. H. Agor (Ed.), *Intuition in Organization*. 23-39. Newbury Park, CA: Sage, 1989.

Velasquez, J. From Affect Programs to Higher Cognitive Emotions: An Emotion-Based Control Approach. *In Proceedings of the Emotion-Based Agent Architecture Workshop at the International Conference on Autonomous Agents*, 1999.

Whehrle, T. The Grounding Problem of Modeling Emotions in Adaptive Artifacts. *Cybernetics and Systems: An International Journal*, 32: 561-580, 2001.

Whehrle, T. and Scherer, K. Toward Computational Modeling of Appraisal Theories. In *Appraisal Processes in Emotion*, Scherer, K., Schorr, A. and Jonhstone, (Eds.), New-York: Oxford University Press, 2001.

APPENDIX C: Publication III

A. Marpaung, S. Brown, and C. L. Lisetti. A Technical Demonstration of Lola, the Robot Entertainer. In *Proceedings of the ACM Multimedia International Conference*, Juan les Pins, France. December 2002

A Technical Demonstration of Lola, the Robot Entertainer

Andreas H. Marpaung
University of Central Florida
4000 Central FL Blvd, BLDG 45
Orlando, FL 32816
1.407.823.3931
marpaung@cs.ucf.edu

Sarah M. Brown
University of Central Florida
4000 Central FL Blvd, BLDG 45
Orlando, FL 32816
1.407.823.3931
sbrown@cs.ucf.edu

Christine L. Lisetti
University of Central Florida
4000 Central FL Blvd, BLDG 45
Orlando, FL 32816
1.407.823.3931
lisetti@cs.ucf.edu

In this paper, we propose a demonstration of Lola - our autonomous intelligent mobile robot produced by ActivMedia.

Lola has the same hardware and software as implemented in her laboratory companion, Cherry [1]. Cherry was designed to be a tour guide and a service robot to help in an office environment, especially for the Computer Science Department at the University of Central Florida. On the other hand, Lola's main role is to entertain and feature some of the state-of-art in multimedia developments integrated onto a robotics platform.

Figure 1. ActivMedia Robot Series

The interface is programmed using ARIA (ActivMedia Robotics Interface for Application) - an object-oriented, robot control applications-programming interface for ActivMedia Robotics' mobile robots. ARIA, as well as the entire interface, is written in C++ and provides access to the robot server as well as to the robot's sensors and accessories.

Interface

The interface will be displayed in a Sony Vaio Picture book that is placed on the top of the robot. In the interface, we have created: 1) several buttons that each have their own function, 2) the avatar – created using Haptek's People Putty that represents Lola's facial expression, 3) Lola's voice, and 4) Lola's vision.

Buttons

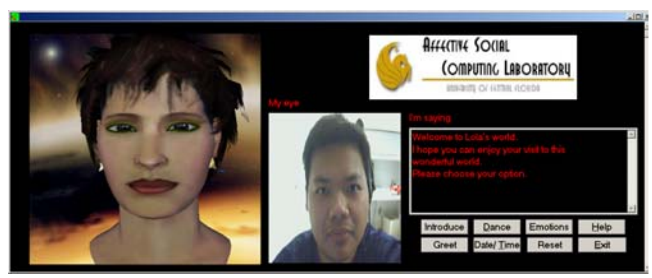


Figure 2. Lola's interface (Neutral Expression)

Currently, we have programmed several buttons with different tasks, which are displayed on Lola's interface as shown in Figure 2 and described below:

a. Introduce – Lola can introduce herself when the users click on this button. In this introduction, she will explain everything about her personality. For example: "Hi! I am Lola. I am an AmigoBot programmed in the Affective Social Computing Laboratory at the University of Central Florida. I like to sing, dance, and do so many things."

b. Dance – Lola can entertain the users by dancing to a variety of tunes and songs available as part of her audio library and played through her speaker system.

c. Emotions – Lola can show some of her emotion through her facial expressions. She is able to show her expression when she falls in love, when she feels angry to someone, etc.

d. Help – Lola can explain how to operate her system properly to the users including the interface features that she has to offer.

e. Greet – Lola can greet a person that she encounters during interactions with groups of people. In order to recognize and greet a certain person, she has to utilize her facial recognition algorithm that compares the captured image with the images in her picture database. Based on the result, she can greet the person if she recognizes him/ her or apologize for not having their image in her current database. She has the capability to distinguish the recognized person based on their status (married/ unmarried) and position (professor/ student) in the department. For example: she greets a professor “Good morning, Dr. Green” and “Hey, dude!” to a student. To unmarried staff, she greets “Good afternoon, Ms. James” instead of “Good afternoon, Mrs. James”.

f. Date/ Time – Lola can tell the user current time and date. With this feature, we are able to greet “Good morning” at 8 AM instead of “Good night”.

g. Reset – Lola can use this button when she reaches the highest state of emotions to downgrade her emotion back to neutral.

h. Exit – Lola will exit the system when the user clicks this option.

Avatar

Lola can express her internal emotional state through her avatar. Currently, she has four different emotions – neutral, frustrated, discouraged, and angry. She begins with neutral and expresses the other three emotions sequentially if the users make her to perform repetitive tasks. For example:

(1) She feels frustrated if users want her to introduce herself more than twice (Figure 3).
(2) After introducing herself more than four times, she feels discouraged because she thinks that the users never listen to her explanation (Figure 4).
(3) When she introduces herself for the sixth time, she reaches her highest state and becomes angry. When she reaches this state, she refuses to allow the users to make her perform a task. She will continue to prompt the user to hit the “Reset” button (Figure 5).

Figure 3. Facial Expression for Frustration

Voice

In order to communicate with the users, Lola can speak through the avatar. We also facilitate a text box that is able to print every word that comes from Lola’s mouth. The text box helps social interaction in noisy environments, where it may be hard for the users to understand her, and for the hard of hearing.

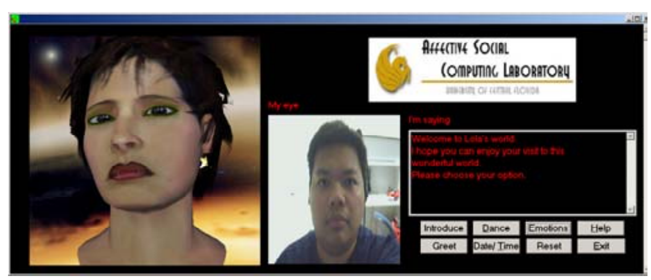


Figure 4. Facial Expression for Discouragement

Figure 5. Facial Expression for Angry

Vision

To be able to capture an image, we add a camera to Lola. In the interface, we display Lola's view of the world. By having this view, the users will be able to position him/ herself in the small rectangle to raise the performance rate of the face recognition algorithm.

A preliminary demonstration has created social interaction between the users and Lola. In the future, we want to raise the level of social interaction by enhancing the ways in which people will be able to interact with her.

References

[1] ActivMedia. www.activmedia.com

[2] Haptek. www.haptek.com

[3] Lisetti, Christine L., Brown, Sarah, Marpaung, Andreas,. Cherry, The Little Red Robot…. with a job…. And her own Personality. *AAAI Fall Symposium*, November 2002

Pictures

Lola with the camera mounted to her

Lola with the picture book on the top of her

LIST OF REFERENCES

Activmedia, Inc. www.activmedia.com. 2002

Anderson, Richard C. "The notion of schemata and the educational enterprise: General discussion of the conference." In Anderson, R. C., Spiro, and Montague 1984.

Arbib, M. “Schema Theory,” *The Handbook of Brain Theory and Neural Networks*, Michael A. Arbib, ed. MIT Press. pp. 830-834.

Arkin, R.C. and Murphy, R.R. “Autonomous Navigation in a Manufacturing Environment,”. *IEEE Transactions on Robotics and Automation.* vol. 6. no. 4. pp. 445-454 August, 1990 .

Breazeal, C. “The framework”, http://www.ai.mit.edu/projects/sociable/kismet.html. 2002.

Breazeal, C. and Scassellati, B. "Infant-like Social Interactions Between a Robot and a Human Caretaker". To appear in *Special issue of Adaptive Behavior on Simulation Models of Social Agents*, guest editor Kerstin Dautenhahn. 2000.

Breazeal, C. “Emotion and sociable humanoid robots”. *International Journal of Human Computer Studies.* Vol. 59. pg. 119 – 155. 2003.

Brown, S. Lisetti, C. Marpaung, A. Cherry, the Little Red Robot with a Mission and a Personality. In Working Notes of *the AAAI Fall Symposium Series on Human-Robot Interaction.* Menlo Park. CA: AAAI Press, 2002.

Damasio, A.R. Descartes’ Error: Emotion, Reason, and the Human Brain. Gosset/ Putnam Press. New York. 1994.

Draper, B., Collins, H., Brolio, A. Hanson and E. Riseman. “The Schema System,” *International Journal of Computer Vision,* vol. 2, pp. 209-250. 1989.

Ekman, P and Friesen, W. "The Facial Action Coding System". Consulting Psychologist Press, San Francisco, CA. 1978.

El-Nasr, Magy Seif. Yen, John. Ioerger, Thomas. FLAME - A Fuzzy Logic Adaptive Model of Emotions, *Automous Agents and Multi-agent Systems,* 3, 219-257. 2000.

Iberall, T. and Lyons, D., “Towards Perceptual Robotics,” proceedings of *1984 IEEE International Conference on Systems, Man and Cybernetics,* Halifax, NS, Canada, Oct. 10-12. pp. 147-157. 1984.

Intel, Inc. www.intel.com. 2002.

Isen, A.M., Daubman, K.A., and Nowicki, G.P. Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology,* 52(6), p. 1122- 1131. 1987.

Isen, A.M. Positive Affect and Decision Making, in Handbook of Emotions, Eds. M. Lewis and J. Haviland, Guilford Press. New York. 2000.

Leventhal, H. "A perceptual-motor processing model of emotion." In P. Pilner, K. Blankenstein, & I.M. Spigel (Eds.). *Perception of emotion in self and others. Vol. 5* (pp. 1 – 46). New York: Plenum. 1979.

Leventhal, H. "Toward a comprehensive theory of emotion." In L. Berkowitz (Ed.). *Advances in experimental social Psychology. Vol. 13* (pp. 139 – 207). New York: Academic Press, 1980.

Leventhal, H. "A perceptual-motor theory of emotion." In L. Berkowitz (Eds.). *Advances in experimental Social Psychology. Vol. 17* (pp. 117 – 182). New York: Academic Press. 1984.

Leventhal, H. Mosbach, P. "A perceptual-motor theory of emotion." In T.J. Cacioppo & R.E. Petty (Eds.). *Social psychophysiology: A sourcebook* (pp. 353 – 389). New York: Guilford, 1983.

Leventhal, H. Scherer, K. "The relationship of emotion to cognition: A functional approach to a semantic controversy." *Cognition and Emotion.* Vol. 1. No. 1. pp. 3 – 28. 1987.

Lewis, V.E. and Williams, R.N. Mood-congruent vs. mood-state-dependent learning: Implications for a view of emotion. Special issue of the Journal of Social Behavior and Personality. 4, 157-171. 1989.

Lisetti, C. Brown, S. Alvarez, K. and Marpaung, A. "A Social Informatics Approach to Human-Robot Interaction with a Service Social Robot". *IEEE Systems, Men, and Cybernetics*. Special Edition on Human-Robot Interaction. Vol. 34 No. 2. May 2004.

Lisetti, C. Nasoz, F. Lerouge, C. Ozyer, O. and Alvarez, K. Developing Multimodal Intelligent Affective Interfaces for Tele-Home Health Care. International Journal of Human-Computer Studies Special Issue on Applications of Affective Computing in Human-Computer Interaction, Vol. 59 (1-2): 245-255. 2003.

Lisetti, C.L Personality, Affect and Emotion Taxonomy for Socially Intelligent Agents. In Proceedings of the *15th International Florida Artificial Intelligence Research Society Conference* (FLAIRS'02), Menlo Park, CA: AAAI Press, 2002.

Lisetti, C. and Bianchi, N. Modeling Multimodal Expression of User's Affective Subjective Experience. *User Modeling and User-Adapted Interaction, An International Journal,* Vol. 12(1): 49-84. 2002.

Maes, P. and R. A. Brooks, "Robot Insect Societies", *Data Manager Magazine,* May, 1990.

Merriam Webster Inc. Merriam Webster Online Dictionary. http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=emotion. 2004.

Michaud, F., Audet, J., Letourneau, Lussier, Theberge-Turmel, C. Caron, S. Artificial Emotions and Social Robotics. *In proceedings Distributed Autonomous Robotic Systems*, 2000.

Michaud, Francois. "Using Motives and Artificial Emotions for Prolonged Activity of A Group of Autonomous Robots". *In proceedings of the 5th International Conference of Autonomous Agent.* 2001.

Michaud, F., Audet, J. "Using motives and artificial emotion for long-term activity of an autonomous robot", *Proceedings 5th Autonomous Agents Conference*, ACM Press, Montréal Québec, 188-189. 2001a

Michaud, F. Audet, J. Letourneau, Lussier, Theberge-Turmel, C. Caron, S. Experiences with an Autonomous Robot Attending AAAI. *IEEE Intelligent Systems,* September-October. 2001b

Michaud, F., Robichaud, E., Audet, J. "Using motives and artificial emotion for prolonged activity of a group of autonomous robots", *Emotional and Intelligent II: The Tangled Knot of Social Cognition* - AAAI Fall Symposium, Cape Cod Massachussetts, Nov, Technical Report FS-01-02, 85-90. 2001c.

Minsky, M. The Society of Mind. New York: Simon & Schuster. 1986.

Murphy, R. R. Introduction to AI Robotics. Cambridge, MA: MIT Press, 2000.

Murphy, R. Lisetti, C. Tardiff, R. Irish, L. Gage, A. "Emotion-Based Control of Cooperating Heterogeneous Mobile Robots", *IEEE Transactions on Robotics and Automation,* Vol. 18, No. 5, October 2002.

Nasoz, F. O. Ozyer, C. L. Lisetti, and N. Finkelstein (2002). Multimodal Affective Driver Interfaces for Future Cars. In Proceedings of the ACM Multimedia International Conference 2002. Juan les Pins, France, December 2002.

Newell, A. & Simon, H. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall.

N. J. Nilsson, ``Shakey the Robot,'' Technical Report 323, SRI International, Artificial Intelligence Center, 1984.

Orange Micro, Inc. http://www.orangemicro.com/ibot.html. 2002.

Ortony, A. On Making Believable Emotional Agents Believable . In Emotions in humans and artifacts, R. Trappl & P. Petta P. (eds.), MIT Press, Cambridge, 2001.

Picard, Rosalind W. Affective Computing, Cambridge, Mass.: MIT Press, 1997.

Picard, Rosalind W. (2001). What does it mean for a computer to "have" emotions? . Emotions in Humans and Artifacts, Eds. R. Trappl, P. Petta and S. Payr. Cambridge, Mass: MIT Press, Cambridge, 2003.

Scherer, K. "Emotion as a multicomponent process: A model and some cross-cultural data." In P. Shaver (Ed.), *Review of personality and social psychology.* Vol. 5. *Emotions, relationships and health* (pp. 37 – 63). Beverly Hills, CA: Sage. 1984.

Scherer, K. "Vocal affect expression: A review and a model for future research." *Psychological Bulletin.* 99. pp. 143 – 165. 1986.

Scherer, K. "Criteria for emotion-antecedent appraisal: A review". In V. Hamilton, et al (Eds.), *Cognitive Perspectives on Emotion and Motivation* (pp. 89 – 126). Kluwer Academic Publishers. Netherlands. 1988.

Smith, C. A., Ellsworth, P.C. "Patterns of cognitive appraisal in emotion." *Journal of Personality and Social Psychology.* 48. pp. 813 – 838. 1985.

Simmons, R. Goldberg, D. Goode, A. et al. GRACE An Autonomous Robot for the AAAI Robot Challenge. *AI Magazine.* Vol. 24. No. 2. p. 51 – 72. 2003.

Stork, David G., Hal's Legacy 2001's computer as dream and reality. Massachusetts Institute of Technology Press. Massachusetts. 1997.

Sony Inc. www.sony.com. 2002.

Tardiff, Russel. "Emotion Augmented Control for Heterogeneous Mobile Robot Teams". Master's thesis. University of South Florida. 2002.

Velasquez, Juan. "Cathexis, A Computational Model for the Generation of Emotions and their Influence in the Behavior of Autonomous Agents". Master's thesis. Massachusetts Institute of Technology. 1996.

Velasquez, Juan. "Modeling Emotions and Other Motivations in Synthetic Agents". *In proceedings of the 14th National Conference of AAAI*. 1997.

Velasquez, Juan. "Modeling Emotion-Based Decision Making". *In proceedings of AAAI Fall Symposium Emotional & Intelligent: The Tangled Knot of Cognition.* 1998.

Weiner, B., Russell, D., Lerman, D. "The cognition-emotion process in achievement-related contexts. *Journal of Personality and Social Psychology,* 37, pp. 1211 – 1220. 1979.