In *Proceedings of 29th Annual German Conference on Artificial Intelligence*, June 2006, Bremen, Germany.

BDI+E Framework: An Affective Cognitive Modeling for Autonomous Agents Based on Scherer's Emotion Theory

Christine L. Lisetti♀
Department of Multimedia Communications
Institut Eurecom
Sophia Antipolis, France
lisetti@eurecom.fr

Andreas Marpaung
Department of Electrical Engineering and Computer Science
University of Central Florida
Orlando, Florida, USA
marpaung@cs.ucf.edu

Abstract

In this article, we propose the design of sensory motor level as part of a three-layered agent architecture inspired from the Multilevel Process Theory of Emotion (Leventhal 1979, 1980; Leventhal and Scherer, 1987). Our project aims at modeling emotions on an autonomous embodied agent, Petra, a more robust robot than our previous prototype - Cherry. Our robot 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. The sensory motor level receives and processes inputs and produces emotion-like states without any further willful planning or learning. We describe: (1) the psychological theory of emotion which inspired our design, (2) our proposed agent architecture, (3) the needed hardware additions that we implemented on the commercialized ActivMedia's robot, (4) Petra's multi-modal interface designed especially to engage humans in natural (and hopefully pleasant) social interaction, and finally (5) our future research efforts.

NOTE: We only became recently aware of the KI CFP deadline and would like to inform the reviewers that our paper is not complete and that we will continue to improve on it in terms of formatting (along Springer LNAI guidelines) and content (as we simply did not have time to write it fully but the work is already performed). while we are waiting for the reviews back.

Introduction

Robotic agents have been of great interest for many Artificial Intelligence researchers for several decades. This field has produced many applications in many different fields, i.e., entertainment (Sony Aibo) and Urban Search and Rescue (USAR) (Casper, 2002; Casper and Murphy, 2002) with many different techniques – behavior-based (Brooks, 1989; Arkin, 1998), sensor fusion (Murphy, 1996a, 1996b, 1998, 2000), and vision (Horswill, 1993). Furthermore, since a relatively new field arose, Affective Computing, new interests of modeling emotion have emerged with the field of Artificial Intelligence and Robotic agents. Picard (1997) defines this new field as computing that "relates to, arises from, or deliberately influences emotions".

Along with the growth of the Affective Computing, researchers have been trying to model emotion in intelligent agents for various applications. Fuzzy Logic Adaptive Model of Emotions (FLAME) is an example of emotion modeling in a non-robot domain (El-Nasr, 2000). The model was implemented using fuzzy logic that combines the affected goals, degree of impact, desirability level, and the goals' importance with inductive learning algorithm in an animal simulation – PETEEI (A Pet with Evolving Emotional Intelligence). Implementations on robots have also been explored. Kismet (Breazeal and Scassellati, 2000; Breazeal, 2003), Graduate Student Attending Conference (GRACE) (Simmons et al, 2003), Cathexis at Virtual Yuppy (Velasquez, 1996, 1998), Leguin and Butler (Murphy et al, 2002) are several robotic agents that include emotion in their designs. Kismet, which was built by the perception, motivation, attention, behavior, and motor systems, interacts with its caretakers by perceiving a variety of inputs from its visual and auditory channels and giving feedback to them through its gaze direction, facial expression, body posture, and vocal babbles (Breazeal & Scassellati, 2000). GRACE, an autonomous agent's entry at the 2002 AAAI Robot Challenge, has an expressive face on the screen and the sensors that include a microphone, touch sensors, infrared sensors, sonar sensors, a scanning laser range finder, a stereo camera head on a pan-tilt unit, and a single color camera with pan-tilt zoom capability. With these sensors, GRACE succeeded in completing

♀Part of this work was accomplished while the author was at the University of Central Florida.

her tasks to move from its starting point to the registration counter and then to the conference hall where it gave the speech about itself with fewer human's involvements.

We, in particular, are extending the architecture model proposed by Murphy et al. (2002) for their two waitering robots Leguin and Butler, an entry at the 2000 AAAI Mobile Robot Competition's *Hors D'Oeuvres, Anyone?* event which has won the *Nils Nilsson Award*. Part of the design was an Emotion State Generator (ESG), which is expanded in our model to process the inputs in more detail that will include willful planning and learning.

We focus on the study of *social expertise* for artificial agents in terms of:

1. internal motivational goal-based activities, and
2. external communicative behavior

As shown in Figure 1, we are focusing on the Socially Intelligent Agent architecture (within the red circle) of the **Multimodal Affective User Interface (MAUI)** paradigm proposed and developed earlier (Lisetti, 2002; Lisetti and Nasoz, 2004).

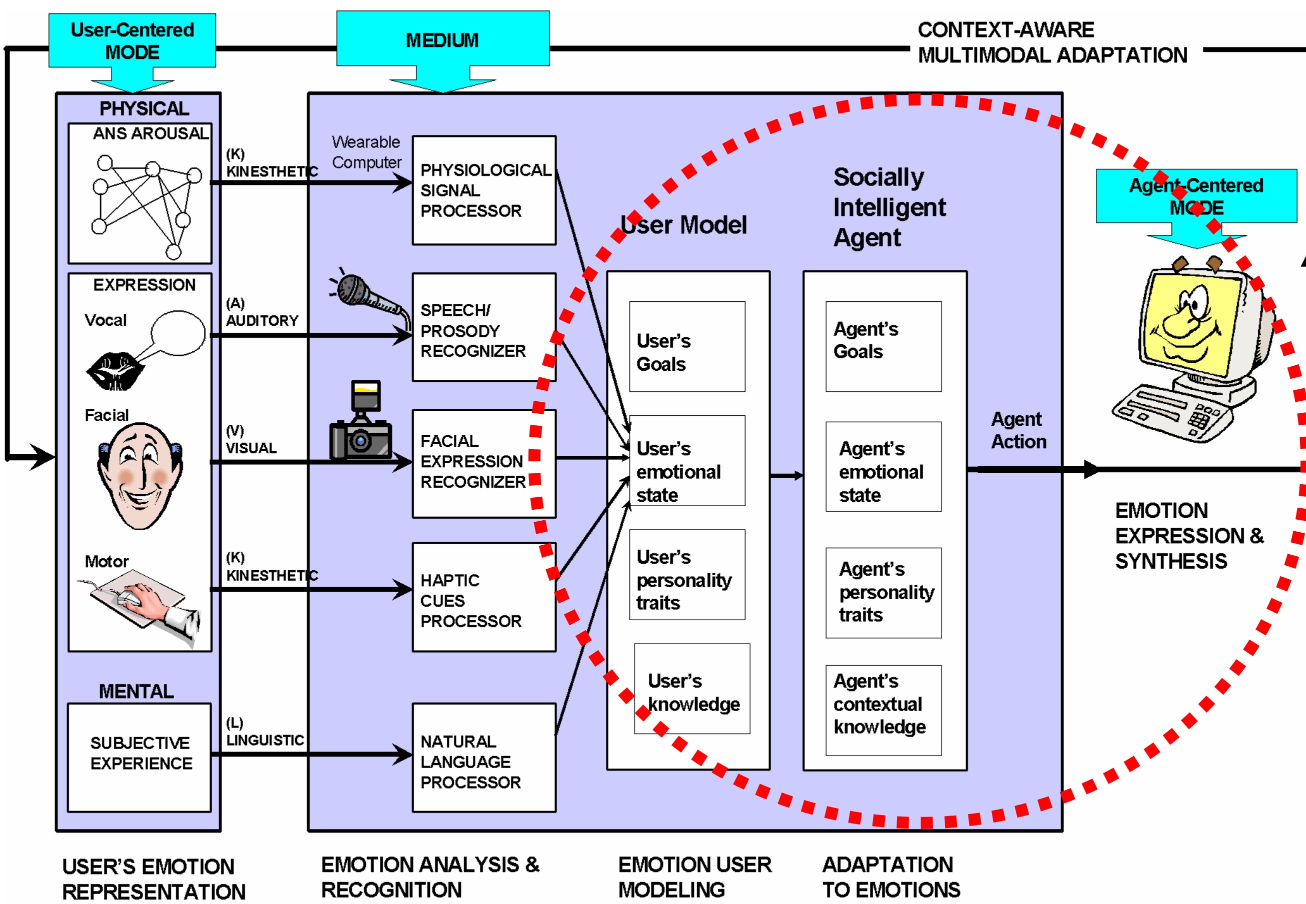

Figure 1: Overall MAUI Paradigm for Multimodal Affective User Interfaces from (Lisetti and Nasoz, 2004)

We propose a psychologically-grounded architecture based on Scherer's affective-cognitive theory of emotions, to be used for the development of artificial agents with diverse forms of embodiment such as vocal robots, graphical animated avatars, avatar-based interface on mobile robotic platform, anthropomorphic robotic platforms (see Section on embodiment, software and hardware).

Multilevel Process Theory of Emotion

With recent advances in Psychology, many researchers have proposed the mechanisms of producing emotions in humans. One of the theories of particular interest to us is the *Multilevel Process Theory of Emotion* (Leventhal 1979, 1980, Leventhal and Scherer, 1987), which we chose to inspire the design and the implementation of the Emotion State Generator (ESG) in our commercially available autonomous robot PeopleBot (ActivMedia, 2002). Figure 1 shows the ESG three-layered architecture we use for generating emotion-like states for our autonomous agents.

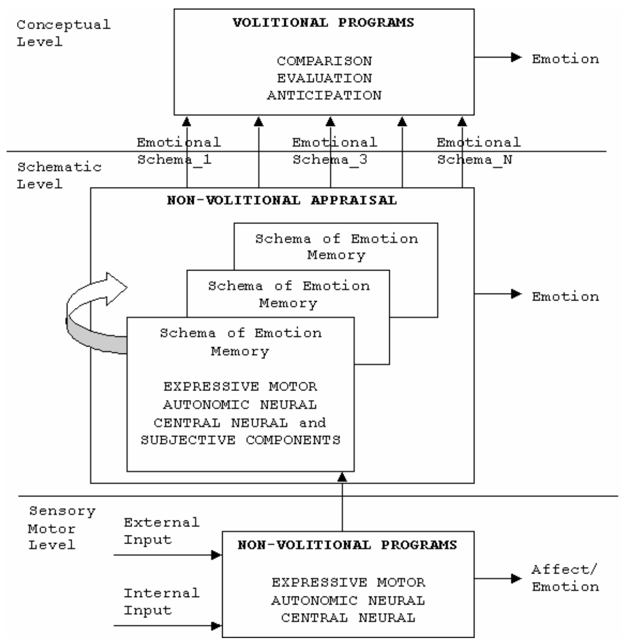

Figure 2: Emotion State Generator (ESG)

Indeed, the Multilevel Process Theory of Emotion postulates that the experience of emotion is a product of an underlying constructive process that is also responsible for overt emotional behavior. It also describes that emotions are constructed from a hierarchical multi-component processing system. In short (Leventhal, 1980):

a. *Sensory motor level* – generates the primary emotion in response to the basic stimulus features in a non-deliberative manner;
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 more closely to social labeling processes.

Sensory Motor Level

The *sensory motor or expressive motor level* is the basic processor of emotional behavior and experience that 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.

Because there is no involvement of the willful planning and learning processes, 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). In this project, we are only modeling: happy, surprise, fear, sad and angry.

We briefly describe the schematic and conceptual levels for completeness sake, but we are currently focusing our design on the sensory motor level.

Schematic Level

The *schematic level* integrates sensory-motor processes with prototypes or schemata of emotional situations in order 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.

Schemata - organized representations of other more elementary codes - are built during 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 that is 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 of as codes, complex categorical units, a network of memory nodes, or perhaps as memory columns that are conceptualized.

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, emotion behavior also has a longer lifetime.

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 of emotional memories with certain concepts that will enable the humans to draw conclusions about their feelings to certain events. By comparing and abstracting information from these schemata with conceptual components – verbal and performance component - the humans can reason, regulate ongoing sequences of behavior, direct attention and generate specific responses to certain events.

The *verbal components* are not only representing the feelings themselves but they are also communicating the emotional experiences to the subject (who can also choose to talk about his/her subjective experience). 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, components can be more sophisticated through active participation of the agent. When performance codes are present, for example, the volitional system can swiftly generate a sequence of voluntary responses to match spontaneous expressive outputs from the schematic system. This volitional system can anticipate emotional behaviors through self-instruction.

Stimulus Evaluation Checks (SECs)

In order to produce emotion for each level, many researchers have hypothesized that specific emotions are triggered through a series of stimulus evaluation checks (SECs) (Scherer, 1984; Scherer, 1986; Weiner, Russell, and Lerman, 1979; Smith and Ellsworth, 1985). Inspired by (Lisetti and Bianchi, 2002), we link the SECs system that performs the emotion components' check in the Affective Knowledge of Representation (AKR) that produces a schema of emotion. This schema can be associated with a certain event and emotion and be part of the schema memory for further use. In AKR, each emotion has many components, e.g., valence, intensity, focality, agency, modifiability, action tendency, and causal chains.

Valence: *positive/ negative:* is used to describe the pleasant or unpleasant dimension of an affective state. **Intensity**: *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.

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).

Agency: *self/ other:* is used to indicate who was responsible for the emotion, the agent itself *self*, or someone else *other.*

Modifiability: *high/ medium/ low/ none:* is used to refer to duration and time perspective, or to the judgment that a course of events is capable of changing.

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.

Causal chain: identifies the causation of a stimulus event associated with the emotion. For example, 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.

Affective-Cognitive Architecture based on the Multilevel Process Theory of Emotion

In terms of architectures for autonomous agents and robots, the multi-level theory of emotions gets translated into the figure 3 below, of which we have implemented various levels and different types of embodiement forms, as we will describe later.

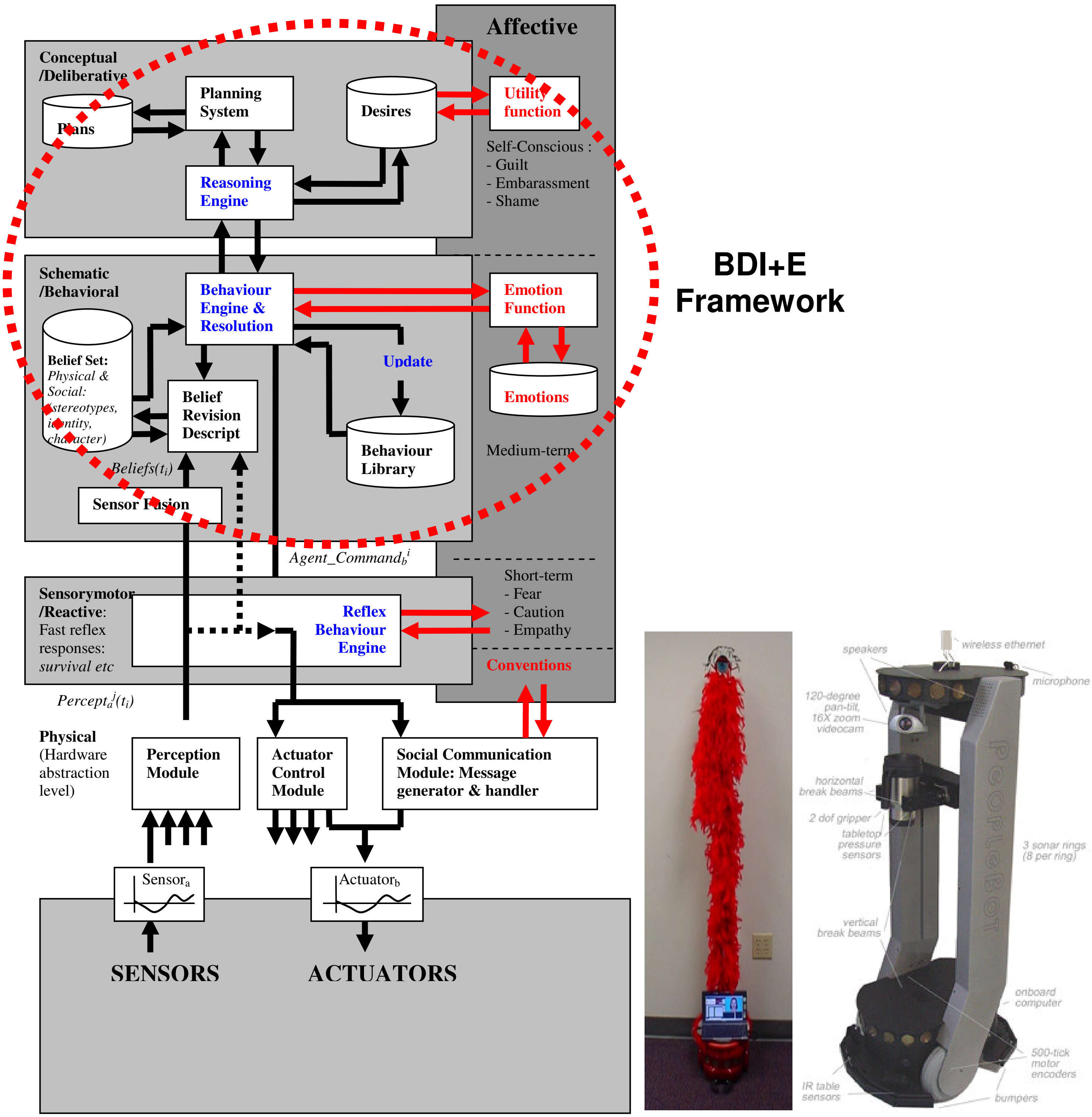

Figure 3: Affective-Cognitive Three-Layered Architecture

Functionalities of our robot

Our robot, Petra, has the same tasks as Cherry (Brown, Lisetti, Marpaung, 2002). Petra is designed so that she can socially interact with humans on a daily basis in the office suite environment especially on the second floor of the computer science building at the University of Central Florida. She has a given set of office-tasks to accomplish, from giving tours of our computer science faculty and staff suites to visitors and to engaging them in social interactions. With the sensors that she has (explained below), she is able to roam around the building using her navigational system and recognize someone through her face recognition algorithm.

Our Agent Three-Layered Architecture

Our ESG is being developed in our robot - Petra, an ActivMedia PeopleBot (ActivMedia, 2002) following the overall architecture shown in Figure 4, which uses the ESG architecture shown in Figure 2 and 3 and discussed in the second section. Currently, Petra has three different sensors - twenty-four sonar, a camera for navigation, and a camera for face recognition to be used during navigation and social interaction. After sensing various stimuli from the real world (e.g., walls, floors, doors, faces), these are sent to the perceptual system. We designed the perceptual system as an inexpensive and simple system so that the information abstracted from the outside world has some interpreted meaning for the robot. For every cycle (in our case, it is 1000 mm travel distance), the sensors send the inputs read to the perceptual system and these are then processed by the perceptual system as described below. Afterward, the perceptual system sends its outputs (valid sonar readings, vision-navigation interpretation, and person's name) to the sensory motor level, which triggers certain emotion-like states.

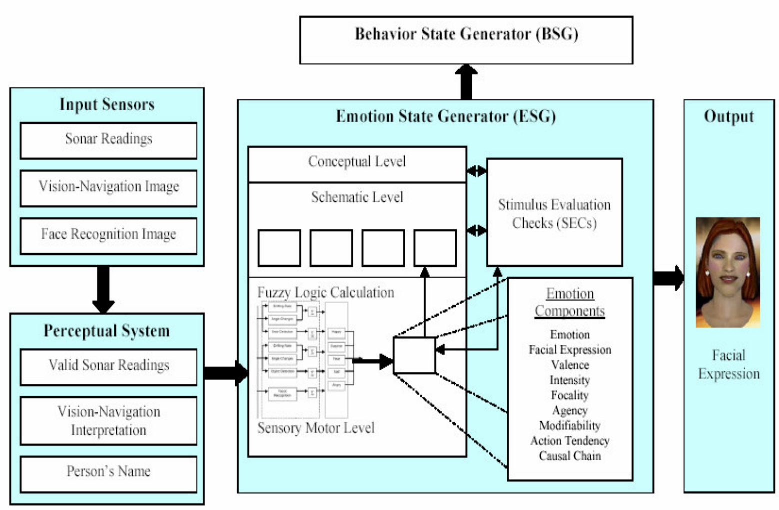

Figure 4: Petra's Detailed Three-Layered Architecture

Sonars

In our design, the robot performs sonar readings every 200 mm, so for 1000 mm, we get five different readings. Out of these five readings, the system extracts the invalid information out and stores only the good ones for further use in the ESG model. The reading is invalid if the sum of the left-most and the right-most sonar readings are extremely more or less than the distance between the aisle (1,500 mm for our case). And vice versa, the reading is valid if the sum of both readings is around 1,500 mm.

Camera for navigation

For every cycle, the camera captures an image and sends it to the vision algorithm. In this algorithm, the image is smoothened and edged by canny edge detector before calculating the vanishing point. In order to calculate the point, in addition to the canny method, we also eliminate the vertical edges and leave the image with the non-vertical ones (edges with some degrees of diagonality). With the edges left, the system can detect the vanishing point by picking up the farthest point in the hall. With this point, represented by the x- and y- coordinate, the system asks the robot to perform course correction, if needed, and uses it as an input for the ESG model. Besides having the capability to center between the aisles of the hallway, the robot is also able to detect some obstacles, i.e, garbage can, boxes, people, etc. When the robot finds the object(s), this detection information is also sent to the ESG model.

Camera for face recognition

The perceptual system receives an input from this camera only when the robot performs the face recognition algorithm. In our current implementation, this algorithm starts when the robot asks someone to stand next to her and captures an image. Along with the FaceIt technology by Identix (Identix, 2002), our algorithm compares the input with her collection of images in her database of 25 images and when any matching is found, she greets that person. The result, recognized or unrecognized along with the person's name (to be used to greet him/her), is also sent as an input to the ESG model. At this level, the other information of the person whose image was captured and recognized (gender, social status, and social interaction value – the degree of her like/dislike toward that person) is not sent to the sensory motor level, but in the future, this information may be needed for the implementation of the schematic and/or the conceptual level where further learning and information processing will be done.

Sensory Motor Level Design

Since the information abstracted from the perceptual system does not go through willful thinking and learning at this level, they may contain some fuzziness to certain degree. Inspired by FLAME (El Nasr, 2002), this level is implemented with the Takagi, Sugeno, and Kang (TSK) fuzzy logic model (Takagi & Sugeno, 1985). Because of its simplicity, it can reduce the number of rules required for this level. Our proposed sensory motor level architecture is shown in Figure 5.

The information received from the perceptual system is then processed further to determine the drifting rates and angle changes which are represented by five fuzzy

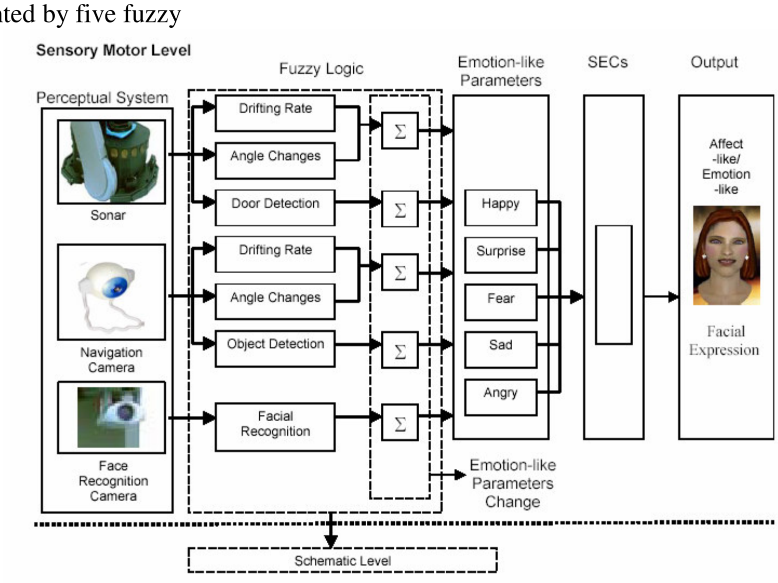

Figure 5: Sensory Motor Level's sub-Architecture

values (small, medium-small, medium, medium-large, and large) and the door detection, the object detection, and the face recognition which are represented by boolean values (found and not-found or recognized and not-recognized). Below are the examples of fuzzy representations of the angle changes calculated from the sonar's valid readings (F_{angle_sonar}). Δ is determined by subtracting the current reading from the previous one.

Small_Angle - Δ angle is between 0^O and 18^O
Medium-small_Angle - Δ angle is between 18^O and 36^O
Medium_Angle - Δ angle is between 36^O and 54^O
Medium-large_Angle - Δ angle is between 54^O and 72^O
Large_Angle - Δ angle is between 72^O and 90^O

Afterward, the processed information (drifting rate, angle changes, door detection, object detection, and face recognition) is calculated with the TSK model that give the emotion-like-parameters-change represented by the numerical value, which will add/subtract the numerical values of the emotion-like-parameters (happy, surprise, fear, sad and angry) based on the OR-mapping shown on Table 1.

$$\sum_{i=0}^{n} \text{Fuzzy\_function}(i) \quad \ldots\ldots\ldots\ldots\ldots (1)$$

where n represents the number of sensors.

Since we have three sensors, equation (1) then can be expanded into:

$$\sum_{i=0}^{n} \text{Fuzzy\_function}(i) = f(\text{drifting-rate, angle-changes}) + g(\text{door-detection}) + f(\text{drifting-rate, angle-changes}) + g(\text{object-detection}) + g(\text{face-recognition}) \ldots\ldots\ldots (2)$$

where f is a fuzzy function that calculates the processed information whose values are represented by the fuzzy values (an example is shown in equation 3 below) and g is a fuzzy function that calculates the processed information whose values are represented by the boolean values.

f(drifting-rate, angle-changes) =

1

$$\frac{\alpha_{drifting\_sonar} * x_{drifting\_sonar} + \alpha_{angle\_sonar} * y_{angle\_sonar}}{\alpha_{drifting\_sonar} + \alpha_{angle\_sonar}} \quad \ldots(3)$$

Parameter	Increased if	Decreased if
Happy	- Small to Medium-small value of the processed information from sonar or vision - Open door - Recognize someone	- Medium to Large value of the processed information from sonar or vision - Closed door - Not recognize someone
Surprise[1]	- 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 - Not recognize someone	- Small to Medium-small value of the processed information from sonar or vision - Open door - Recognize someone
Angry	- Large value of the processed information from sonar or vision (high repetition) - Closed door (repetitively) - Not recognize someone (repetitively)	- Small to Medium-small value of the processed information from sonar or vision - Open door - Recognize someone

Table 1 Mapping of the emotions' parameter changes

where $x_{drifting\_sonar}$ (ε $F_{drifting\_sonar}$) is the fuzzy value's output for drifting rate detected by sonar, y_{angle_sonar} (ε F_{angle_sonar}) is the fuzzy value's output for the angle changes detected by sonar, α_{drift_sonar} is the matching degree for drifting rate detected by sonar and α_{drift_angle} is the matching degree for the angle changes detected by sonar.

Based on the OR-mapping from Table 1, each emotion-like parameter is either increased or decreased based on the numerical result of the emotion-like-parameter-change. The highest value of the emotion-like-parameters is chosen as the effected emotion. Otherwise, when there are ties among the emotion-like-parameters, we only choose the highest rank of emotion-like-parameters, among the ties, in the following order: happy, surprise, fear, sad, and angry.

After calculating the final emotion, this level performs SECs that check the emotion components and create a schema of emotion to be stored in the memory. The checkings are done by assigning appropriate values to the emotion components, as described in the SECs section above, based on the pleasantness, importance, relevance, urgency, etc. Table 2 shows a schema when an unexpected moving object suddenly appears in the captured navigation-image, i.e, walking students. For this case, surprise will be chosen as the final emotion, only for the current cycle.

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 farther distance, the valence is negative. And at current cycle, the modifiability is set to its default–medium because she have not performed the obstacle avoidance to change the course event.

Components	**Values**
Emotion	Surprise
Valence	Negative
Intensity	Very High
Focality	Object – walking student
Agency	Other
Modifiability	Medium
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 something bad

Table 2 A surprise schema

After performing SECs, Petra's facial expression also needs to be changed to show her current emotion-like state. For every emotion-like that we are modeling, e.g., happy, surprise, fear, sad, and angry, we have designed their facial expressions based on the Facial Action Coding System (FACS) (Ekman and Friesen, 1978) as shown in Figure 6 (a-e).

To show surprise, when the processed information from sonar or vision is large on the first detection, the weight of this emotion is the highest among all.

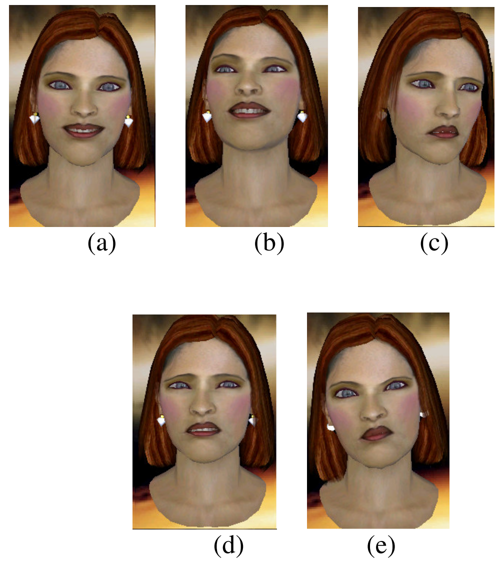

Figure 6: Five facial expressions for the modeled emotions (a) Happy, (b) Surprise, (c) Fear, (d) Sad, (e) Angry

Behavior State Generator (BSG)

A behavior is "a mapping of sensory inputs to a pattern of motor actions, which then are used to achieve a task" (Murphy, 2000). After determining the facial expressions, the processed information is sent to BSG. Through these, she can execute different behaviors depending on the input sources (sonar, camera for navigation, and camera for face recognition). Each behavior state is described below:

1. *INIT*: reset the emotion-like, the progress bars, and the starting position.
2. *STAY_CENTER*: center herself between the aisles to avoid the walls.
3. *AVOID_LEFT_WALL*: move right to avoid the left wall. This behavior is triggered when a course correction, calculated by sonar or vision, is needed.
4. *AVOID_RIGHT_WALL*: move left to avoid the right wall. This behavior is also triggered when course correction is needed.
5. *WAIT*: wait for a period of time when the face recognition algorithm cannot recognize anyone or the door is closed (in order to give another try to avoid any false positive).

Introducing Petra

Petra is a continuation of our previous project – Cherry the AmigoBot, the little red robot (Brown, Lisetti, Marpaung, 2002, Lisetti, Brown, Alvarez, Marpaung, 2004). Cherry, a robot with no formal emotion modeling representation, was developed as a prototype for Petra.

Our main reason of switching from an AmigoBot to the PeopleBot was the size. Realizing that our main goal is to enhance Human-Robot Interaction (HRI), it is better to have a human's height autonomous agent (compared to Cherry). Besides the size, another reason was the weight limitation, the stability, the accuracy, and the on-board computer. With higher weight limitation and better stability and accuracy, we were able to put more hardware that can support our HRI goal, e.g., DesXcape 150 DM, two USB cameras (with 320x240 resolutions), wireless antenna for DesXcape touch screen communication, and laser (for future use). Our hardware designs are shown in Figure 5.

Petra's Interface

Through the on-board computer, we are able to execute the interface that we had created (Figure 7) and display it through the touch screen wirelessly. The interface, the modified version of Cherry's, integrates several components such as the avatar, an anthropomorphic face, a point-and-click map, the emotion changing progress bars, several algorithms (navigation system, vision and obstacle avoidance system, and face recognition system), several help menus, i.e., speech text box, search properties, and start-at-room option, and two live-capture frames. Further information on most of these components can be found in (Brown, Lisetti, Marpaung, 2002). The main differences of Petra's interface from Cherry's are the progress bars, the two video frames, and navigational and vision algorithms. Through these bars, we are able to show the real-time changes of emotion-like state and which emotion-like state(s) is/ are affected by the stimuli accepted. One of the video streams has the same purpose as Cherry's vision for face recognition, and the other one is used for the vision for navigation system. The other two algorithms (navigation and vision) are designed to have a better and smoother navigational system.

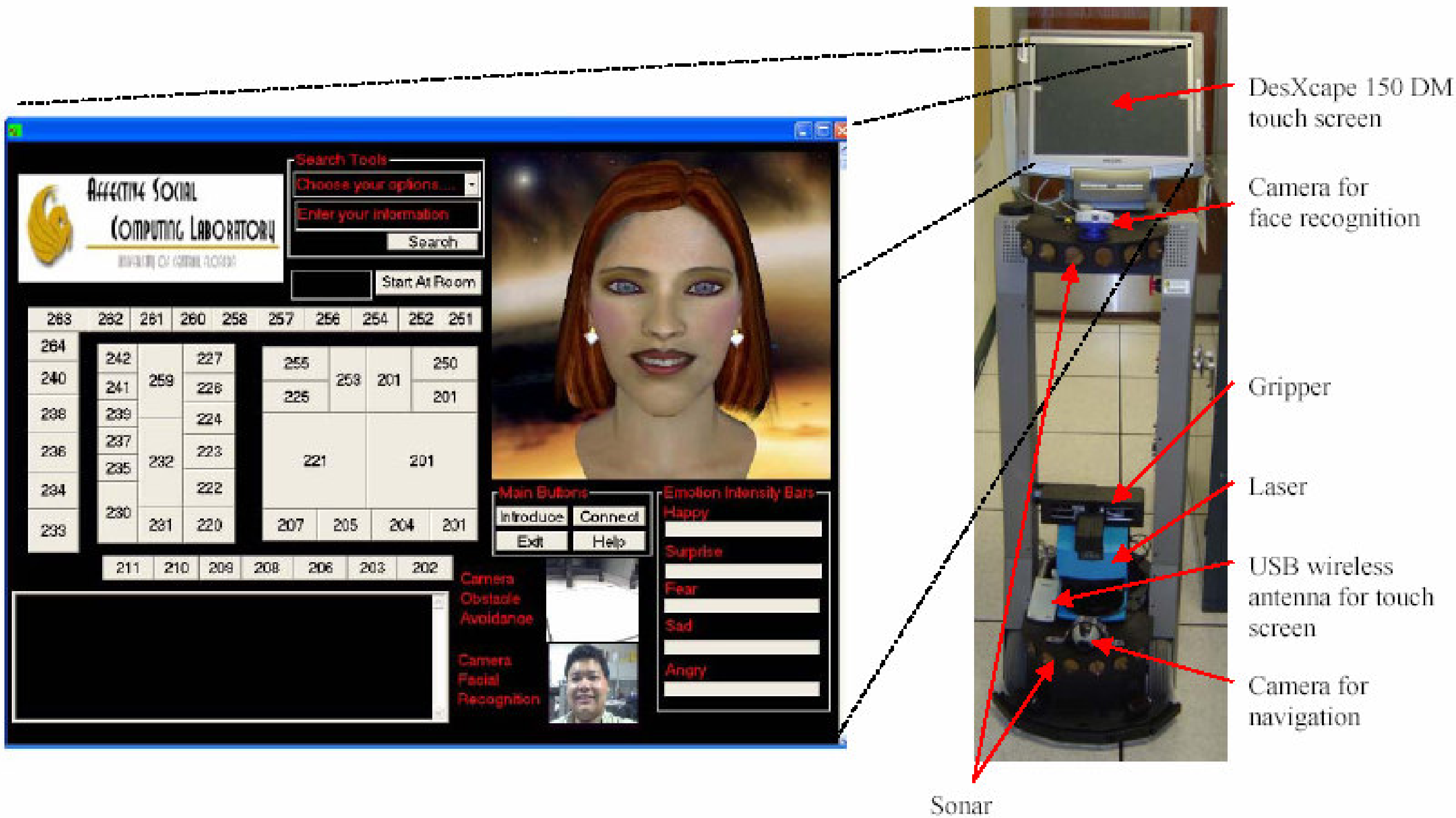

Figure 7: Petra's Complete Interface and Hardware

Our Future Goals:

Implementing Schematic and Conceptual Levels

After implementing the sensory motor level, we want to continue our effort to model emotions to the next levels – schematic and conceptual levels, where further thoughtful reasong, learning, etc happen.

More Personality for Petra

Besides having emotions, we would also plan to create more "personality" for Petra, i.e., sing a song when she is happy, humming as she travels, getting excited when finding someone she loves, etc. In the future, we also can tailor her conversation to the interests of the other person that she speaks to. With these human-like behaviors, we hope that the human-robot interaction can be enhanced so it can be more like human-human interaction.

Acknowledgements

The authors would also like to thank the *Office of Naval Research* for partial funding for this research project. We would also like to thank Eric P. Leger for his help in implementing the navigation and obstacle avoidance algorithms.

References

ActivMedia. www.activmedia.com, 2002.

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

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

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

Brown, S. Lisetti, C. and 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. Cape Cod, MA, November 2002.

Casper, J. "Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center." MS Thesis, Computer Science and Engineering, University of South Florida, April 2002.

Casper, J. and Murphy, R. "Workflow study on human-robot interaction in USAR." In *Proceedings of ICRA 2002*, pp. 1997 – 2003. 2002.

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.

Identix Inc. www.identix.com, 2002.

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. and 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.

Lisetti, C. Brown, S. Alvarez, K. and Marpaung, A. A Social Informatics Approach to Human-Robot Interaction with an Office Service Robot. *IEEE Transactions on Systems, Man, and Cybernetics - Special Issue on Human Robot Interaction,* Vol. 34(2), May 2004.

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

Marpaung, A. "Social Robots with Emotion State Generator Enhancing Human-Robot Interaction (HRI)". Master's thesis. University of Central Florida. in progress 2004.

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

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.

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

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

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