TAME: Time-varying Affective Response for Humanoid Robots

Lilia Moshkina, Sunghyun Park, Ronald C. Arkin, Jamee K. Lee, HyunRyong Jung

Lilia Moshkina, Sunghyun Park, Ronald C. Arkin

Georgia Tech Mobile Robot Laboratory, Atlanta, GA, USA 30332

E-mail: {lilia, ts2883, arkin}@cc.gatech.edu

Jamee K. Lee, HyunRyong Jung

Samsung Advanced Institute of Technology, Kiheung, South Korea

E-mail: {jamee.lee, hyunryong.jung}@samsung.com

Abstract. This paper describes the design of a complex time-varying affective software architecture. It is an expansion of the TAME architecture (Traits, Attitudes, Moods, and Emotions) as applied to humanoid robotics. In particular it is intended to promote effective human-robot interaction by conveying the robot's affective state to the user in an easy-to-interpret manner.

Keywords: *Humanoids, emotions, affective phenomena, robot architectures*

This research is funded under a grant from Samsung Electronics. Portions of this paper are modified from content appearing in Moshkina et al.[1-3].

1 Introduction

With every advance in robotics: visual perception working more precisely, communications becoming more reliable, mobility improving, we are getting ever closer to robots becoming a part of our everyday lives, moving from factories into our homes and workplace. As robots gain more autonomy and start interacting with people not specially trained in robotics, it becomes increasingly important for them to be able to communicate in a way easily understandable to nurses and patients in a hospital setting, or elderly in their homes, or visitors at museums and exhibitions. The development of humanoid robots is certainly a step in the right direction, as they, due to their embodiment, will be more likely to be accepted and used in everyday and long-term situations. However, human-like appearance alone would not guarantee smooth, natural and enjoyable interaction. In particular, humans employ one form of affect or another in almost every step of their lives, and are quite capable of deciphering others' affective nonverbal behavior [4]. Given that people also tend to treat computers as social actors [5] and thus expect them, even if subconsciously, to behave in a socially appropriate manner, we believe that endowing humanoids with affective capabilities would be advantageous for the successful human-robot interaction.

Our research group has had extensive prior experience implementing motivational and affective phenomena in robotic systems. Some of our earlier research included: adding motivational behaviors to a robotic Sowbug and a praying mantis; an implementation of an emotional attachment mechanism in simulation and on real robots; an ethologically-inspired architecture for a robotic dog Aibo with included a number of drives and emotions; and developing emotionally-grounded symbols within EGO architecture on a humanoid robot Sony QRIO (see Arkin [6] for a more detailed summary). Based on our prior experience, we are developing an application of cognitive and psychological models of human Traits, Attitudes, Moods, and Emotions (TAME) for use in humanoid robots. These affective states are embedded into an integrated architecture and designed to influence the perception of a user regarding the robot's internal state and the human-robot relationship itself. Recent work by Arkin et al in non-verbal communication [7] and emotional state for the AIBO [8] addressed powerful yet less complex means for accomplishing these tasks. Introducing time-varying affective states that range over multiple time scales spanning from an agent's lifetime to mere seconds with orientation towards specific objects or the world in general provides the power to generate heretofore unobtained richness of affective expression. This paper describes the cognitive and psychological underpinnings of this work in the context of humanoid robots and affective software architecture, and presents the directions being taken in this ongoing project to implement and test it on a small humanoid robot.

2 Related Work

Although most research on humanoids focuses on the physical aspects (e.g., perfecting walking gaits, sensors or appearance), there are some who also explore affective interaction. For example, humanoid Waseda Eye No. 4 Refined [9] combines emotions, moods, and personality. The overall goal of the system is to achieve smooth and effective communication for a humanoid robot. Although many elements of this system are not psychologically or biologically founded, it provides a few interesting mechanisms, such as modeling personality's influence on emotion via a variety of coefficient matrices and using internal-clock activation component in moods. Another example of a robotic system that combines multiple affective aspects is the Roboceptionist [10]. In this system, emotions and moods are generated as a response to the robot's interaction with passer-bys and events in an imaginary story line; attitudes are treated as long-term moods associated with a certain person or thing. Although the Roboceptionist is not a humanoid, it is equipped with a flat-screen monitor displaying an animated character face through which affect is expressed.

Fukuda et al. [11] also include the notions of emotions and moods in their Character Robot Face; emotions are represented as semantic networks, and the combination of currently active emotions is deemed as mood. Two other humanoid robotic head robots, Kismet [12] and MEXI [13] have emotion and drive systems. Kismet is modeled after an infant, and is capable of proto-social responses, including emotional expressions, which are based on its affective state. In MEXI, the Emotion Engine is composed of a set of basic emotions (positive that it strives to achieve and negative it tries to avoid) and homeostatic drives. In ERWIN, yet another socially interactive robot head, five basic emotions are generated through modulation of hormonal-like parameters [14].

Finally, LEGO robot Feelix [15] is capable of expressing a subset of basic emotions elicited through tactile stimulation.

Other work into humanoid affect primarily focuses on producing recognizable emotional facial and bodily expressions, rather than affect generation. In particular, Nakagawa et al. [16] propose a method to control affective nuances by mapping dimensions of valence and arousal onto velocity and extensiveness of motion and body posture; this method was tested to produce subtle affective expressions on two humanoid robots, Robovie-mini R2 (an upper body humanoid) and Robovie-M (a biped). As another example, Hanson Robotics android head "Einstein" [17] is capable of learning and producing a large number of realistic facial expressions based on Ekman's Facial Action Coding System, FACS [18].

Finally, the research into robotic affect on non-humanoid platforms includes: producing emotional expressions based on the circumplex model of affect on a huggable robot Probo by Saldien et al. [19]; display of affect on expressive robotic head EDDIE [20], based on the circumplex model and Ekman's FACS; expression of Extraversion and Introversion on robotic dog AIBO [21]; and a number of other related endeavors.

3 Cognitive Basis of TAME

In comparison with the aforementioned systems, the TAME framework encompasses a wider range of affective phenomena, and provides psychological grounding for each. It has been initially tested on the entertainment robot dog AIBO [3], and its application to humanoid robots is fairly straightforward in principle. Moreover, humanoid platforms provide certain beneficial affordances for the use of the framework. The first one is their expressive potential, exhibited not only in facial and bodily expressions (e.g., a smile, a shoulder shrug, a handshake), but also in a variety of tasks they could perform for which human-like personalities are readily applicable. Another affordance lies in the emphasis on long-term interaction, and the potential for humanoids to act as partners or companions, rather than bystanders; as some of the components of the framework are subtle in expression and would be noticed best over multiple interactions with the same person.

The *TAME* framework itself takes inspiration from a large number of theories and findings from personality, emotion, mood and attitude areas of psychology [22-24, 26-28, 30-33], which are specifically adapted to the needs of enhancing human-robot interaction. As multiple definitions of affective phenomena exist, here we present those we use for our work:

1. *Affect* is an embodied reaction of pleasure or displeasure signifying the goodness or badness of something;
2. *Personality traits* identify the consistent, coherent patterns of behavior and affect that characterize individuals; although not affective per se, they provide a profound influence on generation and application of affective phenomena;
3. *Affective attitudes* are general and enduring, positive or negative, feelings about an object, a person or a issue;
4. *Moods* are a low-activation, slowly-varying diffuse affective state;
5. *Emotions* are a high-activation short-term affective state and provide a fast, flexible response to environmental contingences in the form of relevant stimuli.

3.1 Overview

The Affective Module, the core of TAME, is subdivided into Dispositions and Affective State. Dispositions include personality traits and affective attitudes, and represent a propensity to behave in a certain way. They are more or less persistent, long-lasting, and either slowly changing (attitudes) or permanent (traits) throughout a robot's "life". Affective state consists of emotions and moods, that are more fleeting and transient affects, and manifest as either high-intensity, short-duration peaks (emotions) or slow smooth undulations (moods). Table 1 summarizes the differences in duration and temporal changes of these four components.

Another direction along which these components differ is object specificity: emotions and attitudes appear and change in response to particular stimuli (such as fear in the presence of an attacker or dislike towards an unfriendly person), whereas traits and moods are diffuse and not object-specific – they manifest regardless of the presence or absence of objects. Each component can be positioned in the two-dimensional space defined by duration and specificity [22-24] (Figure 1). Traits and emotions are at the opposite ends of this spectrum: traits are time-invariant and global, whereas emotions are short-term, dynamically changing and focused. Although all the components can be categorized differently and each can have a distinct function and purpose, these

phenomena cannot be regarded as independent, as they strongly influence each other and interweave to create a greater illusion of life.

	Traits	Attitudes	Moods	Emotions
Duration	Life-long	A few days to a few years	A few hours to a few weeks	A few seconds to a few minutes
Change in Time	Time-invariant	Persistent across time; change slowly with the number of times an object of attitude is encountered.	Change cyclically as a variable of underlying environmental and internal influences; any drastic changes are smoothed across previous mood states	Intensity changes in short-term peaks as eliciting stimuli appear, disappear, and change distance; habituation effects describe decay of emotion even in the presence of stimuli.

Table 1. Summary of Time-varying Aspects of TAME Components

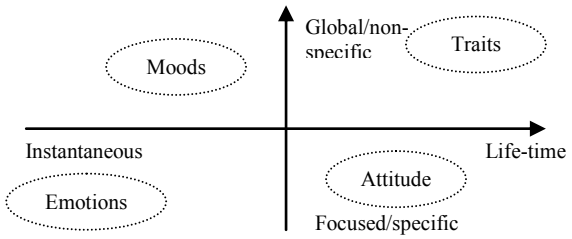

Figure 1: Relative Position of Types of Affect

The Affective Module fits within the behavior-based robotic control paradigm [25] by first processing relevant perceptual input (be it color and distance to certain emotion-eliciting objects or level of light affecting moods) and then directly influencing behavioral parameters of affected low-level behaviors and/or behavior coordination gains as they are composed into behavioral assemblages (Figure 2).

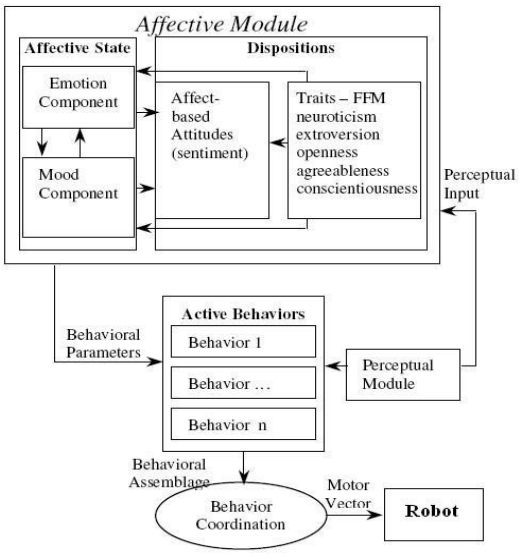

Figure 2: Conceptual View of TAME

3.2 Psychological and Mathematical Foundations

3.2.1 Personality Traits

Personality defines an organism's recurrent patterns of behavior and emotionality. The Five Factor Model of personality traits [26] was chosen as the model for this component for its universality: it is consistent over time, cultures, ages, and even applicable to non-human animals.

To a large extent, traits are inherited or imprinted by early experience, therefore in TAME we treat them as invariable (the exception is that an operator can specify a different personality configuration depending on a task at hand, but it would remain the same during that task). The taxonomy has five broad dimensions, each of which is further subdivided into facets; therefore a robot's personality can be as simple or as complex as desired. Traits provide a two-fold advantage for humanoid robots: first, they serve a predictive purpose, allowing humans to understand and infer the robot's behavior better; second, they allow adaptation to different tasks and environments, where certain trait configurations are better suited for one or another task or environment.

The five global dimensions are *Openness, Conscientiousness, Extraversion*, *Agreeableness* and *Neuroticism*. *Openness* refers to active imagination, preference for variety and curiosity; *Conscientiousness* describes socially desired impulse control that facilitates task- and goal-directed behavior; Extraversion refers to liking people and preferring large groups and gatherings, and also affect positive emotionality; *Agreeableness* is a dimension of interpersonal tendencies, and refers to being sympathetic to others, cooperative and eager to help*;* finally, *Neuroticism* is the general tendency to experience negative affect, such as fear, sadness, anxiety, etc and be more sensitive to signs of danger. Each of them has its own effect on robot behavior: for example, in a humanoid, extraversion could be expressed by keeping a closer distance to the human, frequent smiles, more gestures, etc. This trait would be appropriate for tasks requiring engagement and entertainment from a robot, e.g., a museum guide or a play partner for kids. Another example of a useful trait is Neuroticism: a humanoid can, through correspondingly neurotic behavior, suggest to an accompanying human to pay more attention to potentially dangerous surroundings.

The traits are modeled as vectors of intensity, where intensity refers to the extent to which a trait is represented. In the robot, these intensities: are defined a priori by a human; don't change throughout the robot's "life" (this could be a single run, an interaction with a person, or the robot's entire physical life-span); and are not influenced by any other affective phenomena. We provide a functional mapping from the trait space onto behavioral parameter space as a 2nd degree polynomial, where 3 pairs of corresponding data points are minimum trait/parameter, maximum, and default/average (the values are taken from the normally distributed human psychological data [27]). Traits can have a direct or an inverse influence on particular behaviors and this relationship is defined in a matrix beforehand. Figure 3 presents a 2nd degree polynomial mapping from the trait of Neuroticism onto two behavioral parameters: directly to obstacle avoidance gain (a degree to which an agent should avoid obstacles) and inversely to wander gain (related to exploration).

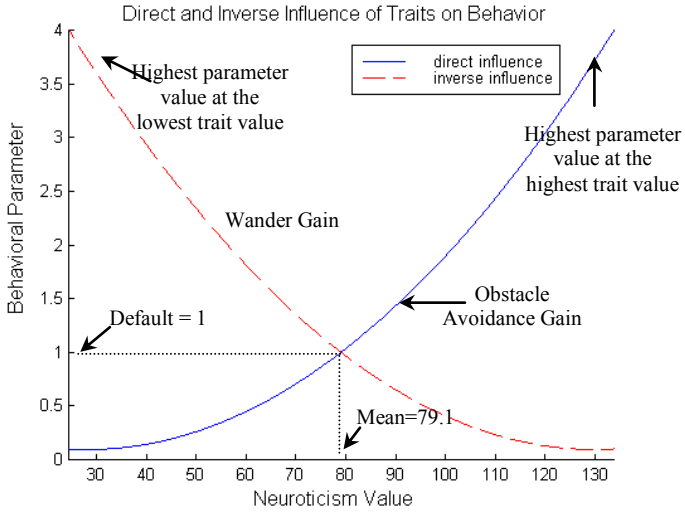

Figure 3: Comparison of Direct and Inverse Influences of Traits on Behavior

In cases where multiple traits affect the same behavior (e.g., Neuroticism may push the robot away from the obstacles while Conscientiousness could make it go closer for a faster route), first a trait/parameter mapping is calculated, according to the chosen function $f_{ij}(p_j)$, where trait i influences behavior j, a polynomial in this case. Then, the results are averaged across all influencing personality traits to produce the final parameter value used thereafter:

$$B_j = \frac{1}{\sum_{i=1}^{N} pb_{ij}} \sum_{i=1}^{N} f_{ij}(p_i) \qquad (1)$$

where B_j is a particular behavioral parameter, $f_{ij}(p_i)$ is the function that maps personality trait p_i to B_j, N is the total number of traits, and $\overrightarrow{pb}$ is personality/behavior dependency matrix; if there is no influence, the result of $f_{ij} = 0$.

As the traits are relatively time-invariant, the obtained trait-based behavior parameters serve as default behaviors for the robot.

3.2.2 Emotions

From an evolutionary point of view, emotions provide a fast, flexible, adaptive response to environmental contingencies. They appear as short-term, high-intensity peaks in response to relevant stimuli (we don't usually live in a constant flux of emotions), and serve a number of functions, of which most applicable for humanoids are communicative, expressive and affiliative, e.g., fear communicates danger and a request for help, while joy in response to a bright smile helps forge trust and camaraderie. The primary, reactive emotions of *fear, anger, disgust, sadness, joy and interest* were chosen, in part because these basic emotions have universal, well-defined facial expressions [28], are straightforwardly elicited, and would be expected, perhaps subconsciously, on a humanoid's face, as appearance does affect expectations. Each emotion's intensity is stored in the emotion intensity matrix $\vec{E} = [E_i]$, where $0 \le E_i \le g_i$, the value E_i represents the intensity of a currently active emotion, 0 signifies the absence of emotion, and g_i is the upper bound for emotion i.

From an emotion generation point of view, Picard [29] suggests a number of properties are desirable in an affective system:

1. **Activation.** Refers to certain stimulus strength below which the emotion is not activated.
2. **Saturation**. Refers to the upper bound of an emotion, after which, regardless of the increasing stimulus strength, the emotion doesn't rise any more.
3. **Response decay**. States that emotions decay naturally over time unless they are re-stimulated.
4. **Linearity**. Emotions can be modeled as linear under certain conditions; due to the properties of activation and saturation, the emotions will approximate linearity only for certain stimulus strength range, and will approach a sigmoid at its edges.

Taking these properties into consideration, the resulting function for emotion generation (based on stimulus strength) resembles a sigmoid, in which the left side corresponds to activation, the right side corresponds to saturation (amplitude), and the middle models the actual response. The eliciting stimulus strength for each emotion is calculated by taking into account a number of object properties, such as its physical properties (size, shape, etc.), its position (distance to the object, its velocity, etc.), and any existing attitude of the agent towards the eliciting stimulus. Then, the base emotion level is calculated as follows:

$$E_{i,base} = \begin{cases} (s_i - a_i)/d - g^{(-a_i)/d_i}, & if \quad a_i \le s_i < (a_i + b_i)/2 \\ g_i - g^{(s_i - a_i)/d_i}, & if \quad s_i \ge (a_i + b_i)/2 \end{cases}$$
$$where \quad b_i = 2 \cdot d_i \cdot \ln(g_i + g^{(-a_i)/d_i})/2 + a_i \qquad (2)$$

where $E_{i,base}$ is the base emotion value for emotion i, s_i is the strength of stimulus eliciting emotion i, a_i is the variable that controls the activation point for emotion i, d_i is the variable that controls the maximum slope for emotion i, g is the amplitude of emotion i, and b_i is the break-point, at which the emotion reverses its rate of growth. Figure 4 presents the resulting curve graphically.

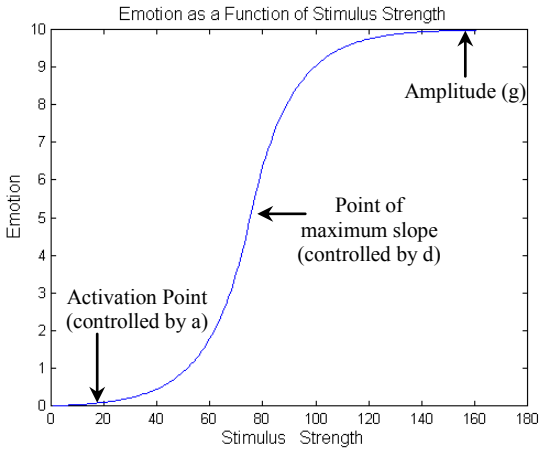

Figure 4: Emotion Generation Based on Stimulus Strength

Emotions are also highly dependent on traits and moods: personality may influence the threshold of eliciting stimulus (activation point), peak (amplitude) and rise time to peak (affecting the slope of the generation curve) [30]; and moods can vary the threshold of experiencing a specific emotion [31]. For example, Extraversion is correlated with positive emotions, therefore a humanoid robot high in this dimension would display more smiles, excited gestures and other expressions of joy. Attitude also has an effect on emotion – the object of like or dislike may serve as a stimulus for emotion generation. A linear mapping from traits to amplitude, activation point, and maximum slope is used to obtain personality influence on emotion generation. For example, the trait of Extraversion provides a direct influence on the amplitude, activation point, and slope of positive emotions (joy and interest), therefore a robot with a higher level of Extraversion will have a stronger positive emotion that will be activated at weaker stimulus strength and will rise faster than that of an introverted robot. Similarly, current mood will influence the activation point, where the negative mood will make it easier for an agent to experience negative emotions, and positive mood – positive emotions. Figure 5 presents combined influence of traits and mood on emotion generation.

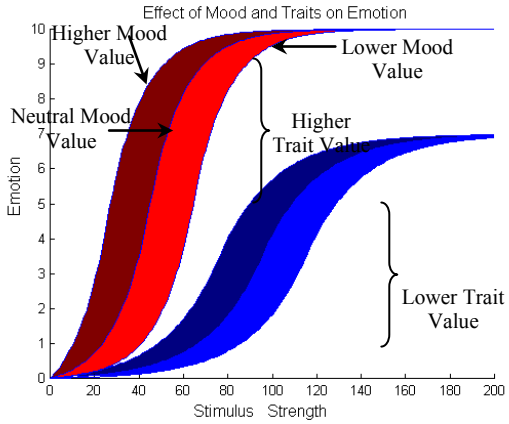

Figure 5: Combined Mood and Trait Influences on Emotion Generation

To account for the short-term duration of emotions and habituation to prolonged stimulus, emotion decay is modeled as a slowly decreasing exponential:

$$E_{i,t,decay} = E_{t,base} - ,^{(t-\ )*d} \qquad (3)$$

where $E_{i,t,decay}$ is the intensity of emotion i at time t, t_o is the time at which emotion is activated (becomes greater than 0), and d is a variable that controls the rate of decay. This ensures that a high-activation emotional state is not maintained beyond the initial episode, and, provided the stimulus doesn't change, the emotion it invoked will eventually dissipate.

Finally, in order to smooth the emotion change in cases of sudden appearance and disappearance of eliciting stimuli, a weighted averaging filter can be used:

$$E_{i,t,filtered} = w_{current} * E_{i,t,decay} + v_{prior} * E_{i,t-\Delta,filtered}) / (w_{current} + v_{prior}) \quad (4)$$

where $E_{i,t,filtered}$ is the final intensity of emotion i at time t after filtering, $w_{current}$ and w_{prior} are weighting variables controlling the relative importance of current and previous emotional states. This filtering function will help to account for short-term lingering emotions even after the eliciting stimulus has disappeared.

Emotions can have a varied impact on behavior, from a subtle slowing to avoid a disgustful object to a drastic flight in response to extreme fear. This effect can be modeled by linear mapping from emotion strength to relevant behavioral parameters, and Figure 6 provides a comparative view across time of stimulus strength (an object appears, comes closer, and then is gone), corresponding emotion activation (after response decay and smoothing), and the Object Avoidance Gain (which causes an avoidance response to Fear); duration is plotted along the x axis, and normalized values for stimulus strength, fear and object avoidance gain along the y axis. If the object continued to be present and unchanged, then Fear would eventually be brought down to 0.

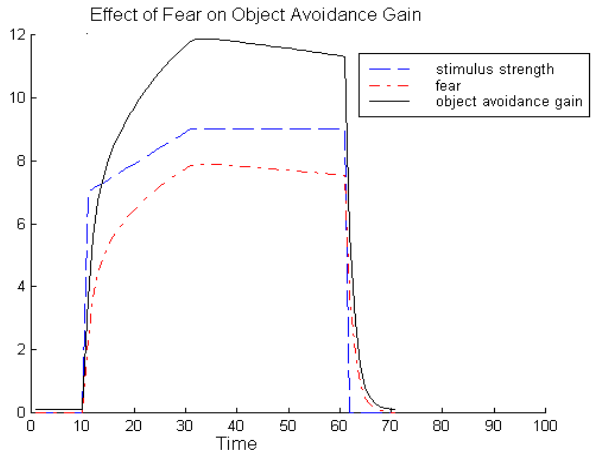

Figure 6: Example of Fear to Object Avoidance Gain Mapping

In a humanoid, display of fear may signal imminent danger to nearby people, and be more persuasive than words alone, in case an evacuation is required. Expressions of disgust, similarly, may alert a human to the presence of some noxious stimulus, which, though not necessarily hazardous, may still be best avoided.

3.2.3 Moods

Unlike emotions, moods represent a global, continuous affective state, cyclically changing and subtle in expression. Mood can be represented along two dimensions, *Positive Affect* and *Negative Affect* [24], where *Negative Affect* refers to the extent to which an individual is presently upset or distressed and describes level of stress and tension, and *Positive Affect* generally refers to one's current level of energy, enthusiasm, and pleasure. The level of arousal for both categories can vary from low to high; a low positive mood value has a negative connotation ("sluggish", "disinterested") and refers to insufficient level of energy, pleasure and enthusiasm, rather than just low. One advantage of providing a humanoid with expression of mood would be to let nearby humans know when the system needs attention, be it a low battery resulting in low energy level, or insufficient amount of light resulting in poor sensor reliability.

There are two broad types of mood change: environmental/external (light or noise level, external temperature, amount of recent interaction, etc.) and internal (e.g., battery level in case of a robot, internal temperature); and short-term situational variables, including emotional episodes. The current base level of mood of a robot is defined as a weighted summation of various external and internal variables. Assuming that the same variables affect both positive and negative moods, strengths of environmental and internal influences can be represented in a matrix $\vec{l} = l_i]$, where $0 < \ _i < \ _i$, where b_i is the hardware-dependent upper bound (e.g., light can only be detected up to a certain level, etc.). The relative weights for each variable are stored in the mood generation matrix $\overrightarrow{mg} = mg_{ij}]$. The values in this matrix are unit conversion factors, to convert the various mood generation variables (which may correspond to raw sensor data) to the same unit, and are

found experimentally for each variable. In addition, negative mg_i stands for inverse influence of the variable on the mood, and positive mg_i stands for direct influence.

According to Set Point theory [24], a certain base level of mood is maintained at all times, and, though events and changes in the environment cause it to fluctuate, it tends to return to the same level over time. As moods are continuous, always present streams of affect, the base mood is continuously generated based on the current environmental and internal influences as follows:

$$m_{base} = \begin{bmatrix} m_{positive} \\ m_{negative} \end{bmatrix} = \sum_{i=1}^{N} \begin{bmatrix} mg_{i,positive} \\ mg_{i,negative} \end{bmatrix} (l_i - l_{i,neutral}) \quad (5)$$

where $\overrightarrow{mg}$ is the mood generation matrix, $\vec{l}$ is the mood generation variable strength matrix, $l_{i,neutral}$ is the set point for mood i, and N is number of mood generation variables. Figure 7 illustrates the effect of an environmental variable on mood. Figure 7 presents an example of the influence of light on mood generation.

Additionally, similarly valenced emotions can affect the corresponding mood intensities additively, and change the existing base mood level in the following manner:

$$\begin{aligned} m_{positive} &= n_{positive,base} + \zeta_k, \ if\ E_k \in \{Interest, Joy\} \\ m_{negative} &= n_{negative,base} + \zeta_k, \ if\ E_k \in \{Fear, Anger, Disgust, Sadness\} \end{aligned} \quad (6)$$

where $m_{positive}$ is the emotion-based intensity of positive mood, $m_{positive}$ is the emotion-based intensity of positive mood, and $\vec{E}$ is the emotion intensity matrix.

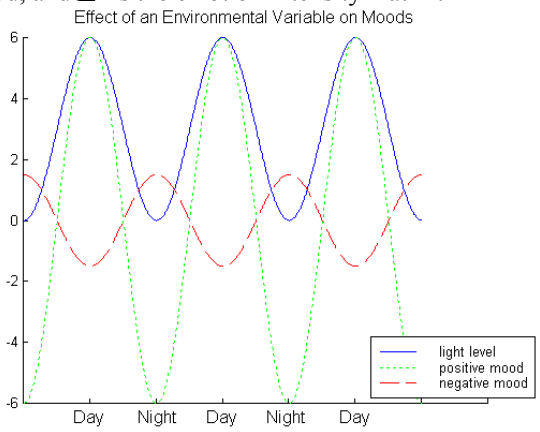

Figure 7**:** Different Effect of Light on Positive and Negative Mood

As mood is a low-activation, slow-varying affective state, sudden changes are smoothed out by taking into consideration prior mood states. Filtering over a longer period of time results in slower and smaller mood changes and helps tone down any drastic spikes due to emotions.

Additionally, in humanoids that are designed for sharing living conditions with humans for a prolonged time, circadian variations in mood may be introduced to provide mood congruency with the human, where user-defined cyclical daily, weekly and seasonal high and low points would be superimposed onto the base mood values.

Moods are mild by definition, and would only produce a small, incremental effect, or a slight bias, on the currently active behaviors. Moods can have a direct or inverse influence on a behavioral parameter. A behavior-mood dependency matrix $\overrightarrow{mb} = [mb_{ij}]$ is defined, where $mb_{ij} \in \{-1, 0, 1\}$ is defined, where –1 corresponds to inverse influence, +1 to direct influence, and 0 to absence of mood influence on behavior. Positive and negative moods may influence the same behavioral parameters, and this influence is treated as additive. As moods are updated continuously, new mood-based values of behavioral gains/parameters replace the existing trait-based values in the following manner:

$$B_{i,mood} = B_{i,trait} + K_i \sum_{j=1}^{N} mb_{ji} \cdot m_j \qquad (7)$$

where $B_{i,mood}$ is the updated behavioral parameter i, mb_{ij} is the mood-behavior dependency matrix value for mood j, m_j is the current value of mood j, N is the total number of mood categories (2), and K is a scaling factor to ensure that the moods produce only incremental effect as opposed to overpowering any of the parameters.

Figure 8 shows an example of incremental effects of moods on behavior. Suppose that mood can bias robot's obstacle avoidance behavior. For example, if visibility is poor, it may be advantageous to stay farther away from obstacles to accommodate sensor error, and vice versa, in good visibility it may be better to concentrate on task performance. Thus, negative mood can bias the obstacle avoidance gain by raising it, and positive mood by lowering it. Neuroticism also affects it by setting the default parameters to be used throughout the life-cycle, and the incremental effect of moods is shown against the space of trait-based defaults (plotted in solid blue center line).

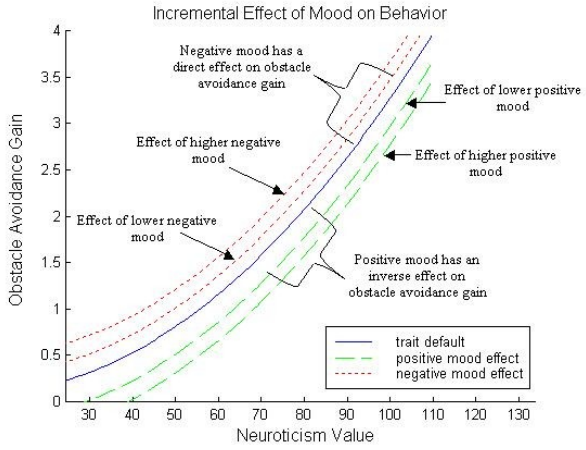

Figure 8: Direct/Inverse Mood Effects on Behavior at Different Neuroticism Values

For human-robot interaction, expressive manifestation of mood can alert a person to favorable or unfavorable changes in the environment or in the robot itself, especially if perception of these changes is based on sensor input not available through human senses. Consider the following scenario. A humanoid is guiding a human inspector through a partially secured search-and-rescue site, when the lights become dim. Although no immediate danger is visible, the robot's negative mood rises, and it displays the signs of anxiety and nervousness; no action per se is warranted yet, but the inspector, picking up the cues from the robot, becomes more alert and ready for action. We are currently using this scenario to formally evaluate, through human subject experiments, the effect of robotic mood display on compliance with and persuasiveness of a robot's request.

3.2.4 Affective Attitudes

From a multitude of definitions of attitudes, the following was adopted as the working definition for TAME: "a general and enduring positive or negative feeling about some person, object or issue" [32]. It stresses relative time-invariance ("enduring"), object/situation specificity, and the role of affect/affective evaluation in the attitude concept. We propose two methods for attitude formation: one is more general and does not require any input from an interacting human, and the other is more experiential, and requires initial human input.

In the general method, attitudes are described by valence and intensity, and are represented as a single value A, ranging from $-\infty$ to ∞, where 0 signifies a neutral (or absence of) attitude, negative values represent increasingly strong negatively-valenced attitude (ranging from a mild dislike to hatred), and positive values refer to increasingly strong positively-valenced attitude (e.g., from a subtle like to adoration). Attitudes are object-specific, and an initial attitude for a particular object (y) would consist of a combination of positive or negative attributes of this object (as a facetious example, a robot may develop a dislike to a man with a mustache), represented as a matrix $\vec{o}_y = o_{iy}]$, where $-\infty \leq o_{iy} < \infty$. Such attributes are not limited to properties of the object only; for example, an emotion invoked by the object and any actions taken by the object

may be considered "attributes". The initial value of the attitude for object *y* ($A_{y,init}$) is calculated as follows:

$$A_{y,init} = \sum_{i=1}^{N} o_{iy} \qquad (8)$$

where $A_{y,init}$ is the newly-formed attitude for object *y*, o_{iy} is an attribute *i* of object *y* that is involved in the attitude formation, and *N* is the number of attributes for object *y*.

Assuming that an initial impression is the strongest, substantial changes in attitude are fairly hard to achieve, therefore any subsequent exposure to the same object would result only in incremental change. This is done by discounting any additional positive or negative object attributes to a certain extent. The updated attitude value for object *y* for n-th encounter ($A_{y,n}$) would then be calculated as follows:

$$A_{y,n} = A_{y,n-} + \lambda \cdot (\sum_{i=}^{N} o_{iy}) \qquad (9)$$

where $A_{y,n-1}$ is the attitude towards object *y* at encounter *n-1*, *n* is the total number of encounters up to date, $\vec{o}_y$ is the matrix of attributes for object *y*, and λ is the discount factor.

Finally, consistent with the finding on mood-congruent judgment, positive mood increases the value of the attitude (a_y) towards an object *y*, and negative mood – decreases it as follows:

$$A_{y,mood} = A_y + \zeta(m_{positive} - n_{negative}) \qquad (10)$$

where $A_{y,mood}$ is mood-enhanced value of agent's attitude towards object *y*, A_y is the original value of agent's attitude towards object *y*, $m_{positive}$ is the current positive mood value, $m_{negative}$ is the current negative mood value, and *K* is a scaling factor to bring moods and attitudes to the same units.

In the experience-based method, robotic attitudes are based on those held by people commonly interacting with robot. In this method, a Case-Based Reasoning approach [33] is used, where an initial set of cases is provided by each interacting human, and the resulting attitudes are expressed through corresponding emotions by the robot. Each case contains a set of object properties (indexed by these properties and a user ID, to differentiate between different people) and corresponding attitudes for a variety of objects. When a robot encounters a new object, the most similar case is retrieved from the case library and applied.

As affective attitudes are closely related to emotions (in fact, some psychologists even describe love and hate as emotions, albeit long-term and persistent), the output of both methods produces stimulus strengths to generate corresponding emotions. Thus, attitudes are not expressed in behavioral changes per se, but rather through the emotions they invoke. In the course of long-term interactions with people, it may be easier for humanoids that share attitudes with their human companions to establish rapport and understanding. For example, a child playing with a robot nanny or tutor may feel greater affinity towards the robot that acknowledges the child's likes and dislikes in toys and games.

3.3 Exploratory Experimental Study

In order to explore the issues of feasibility and potential usefulness of the *TAME* framework, a preliminary exploratory study was conducted (please refer to Moshkina et al [3] for a detailed report). For this between-subject longitudinal human-robot interaction study a partial implementation of Emotion and Trait modules was performed on a robotic dog AIBO. The emotions of Interest, Joy, Anger and Fear were expressed via head, ears and tail positions and movement, a variety of gaits, and LED display; and Extraverted and Agreeable personality was achieved by modifying the percentage of time the robot spent walking around and wagging its tail. The two conditions used in the study were *Emotional* (with the aforementioned affective expressions) and *Control* (without affect). The study was set as "robot as a personal pet and protector scenario", in which 20 people interacted with the robot for a total of four sessions, ranging from 20 to 60 minutes each. For each session, the participants were given one or two new commands to introduce to the robot (7 total: "Stop", "Go Play", "Follow the Ball", "Kick the

Ball", "Follow Me", "Come to Me", and "Sic' em"), and the last session was cumulative. The measures used for this study were: PANAS (mood) questionnaire [35] to assess participants' Negative and Positive mood at the end of each session; Mini-Markers Big-Five personality questionnaire [34] to assess the subjects' personality in the beginning and the robot's personality at the end of the study; and a post questionnaire, to evaluate ease of use and pleasantness of interaction. The post questionnaire consisted of six 5-point Likert scale questions with three subquestions, with "Strongly Agree" anchored at 5, and "Strongly Disagree" anchored at 1. The questions were as follows:

1) It was easy to get the robotic dog perform the commands;
2) It was easy to understand whether the robotic dog was performing the command or not;
3) The robotic dog showed emotional expressions;
4) The robotic dog had a personality of its own;
5) With every session, I was getting more attached to the dog;
6) Overall, I enjoyed the interaction with the robotic dog.

If the participants answered "Agree" or "Strongly Agree" to questions 3 or 4, they were also asked to answer questions 3a,b and 4a, respectively. The subquestions were as follows:

3a) Emotional expressions exhibited by the dog made the interaction more enjoyable;
3b) Emotional expressions exhibited by the dog made the interaction easier;
4a) I enjoyed interacting with the robot, partly because it possessed some personality.

A total of 20 people participated in the study, 10 males and 10 females, distributed equally between the two conditions. The subjects were recruited via flyers posted on and around the Georgia Institute of Technology campus, and they varied widely in the demographics according to age (from between 20 and 30 to over 50 years old), their educational level and backgrounds (from High School diploma to working on a Ph.D., with majority having either a Bachelor's or Master's degrees), and computer experience. Most of the participants had owned pets at some point in their lives (18 out of 20), and had either no or very limited robot interaction experience (only 2 out of 20 had interacted with mobile or entertainment robots prior to the study). A number of interesting and encouraging observations were made in this study, as reflected in the results of 1-tailed Independent Samples T-tests and Pearson Correlations (unless specified otherwise).

First, those participants who believed that the robot displayed emotions (5 out of 10 in the *Control* condition, and 8 out of 10 in the *Emotional* condition) and/or personality (6 out of 10 in both conditions) also believed that these features made their interaction more pleasant: the average answer for question 3a was 4.46, and for question 4a was 4.25. There was also a number of significant correlations between questionnaire responses regarding the pleasantness of the interaction: 1) the response to question 4 (robotic dog had a personality) had a significant positive correlation ($r = .66$, $p=0.02$) with the response to question 5 (participant got more attached to the dog); 2) the response to question 3a (emotional expressions made the interaction more enjoyable) had a significant positive correlation with response to questions 5 ($r=617$, $p=.025$) and 6 ($r=.749$, $p=.003$, "enjoyed interaction with the robotic dog"). Although there was no significant statistical difference between the two conditions in emotion display recognition (question 3), possibly due to small sample size ($M_{non\text{-}emotional}=2.7$, $M_{emotional}=3.6$, $F=.693$, $p<0.088$, $eta^2 =.1$), the perceptions of the robot's emotionality and personality seem to make the interaction more enjoyable and result in greater attachment. This was encouraging, as it suggested that people value expression of emotion and personality in their interaction with an autonomous entertainment robot.

Second, we observed a reduced level of Negative Mood in the subjects in the Emotional condition ($M_{non\text{-}emotional}=13.9$, $M_{emotional}=12.125$, $F=6.462$, $p<0.048$, $eta^2 =.146$), which suggests that affective behavior contributes to the quality of interaction. Additionally, a significant positive correlation ($r=.598$, $p=.007$) between average Positive Mood and the response to question 4 (robot displayed personality) was observed, thus providing a link between perceived robotic personality and users' improved mood.

Finally, women were found to be more attuned to emotional expressions and more ready to attribute emotions to the robot than men: 2-factor ANOVA on Gender and Emotionality resulted in a significant main effect of Gender on the answer to question 3: display of emotions ($M_{female}=3.8$, $M_{male}=2.5$,$F =4.829$, $p<0.043$, partial $eta^2 =.232$). This should be taken into consideration for systems adapted to groups with gender-biased compositions. Other lessons learned from this study included the following: 1) the physical platform affects the perception of emotionality; 2) for between-subject experiments, affective expressions may need to be exaggerated to be consciously acknowledged; 3) there is a great need for suitable repeatable metrics to evaluate usefulness of robotic affect via human-robot interaction studies.

4 Software Architecture

In this section, we introduce the overall software architecture for our affective system, also referred to as *TAME Module* henceforth (Figure 9). The system is designed to be a stand-alone process to achieve maximum platform-independence. With an interface to connect to the system's *TAME Communication Manager* (to supply sensory data), and appropriate configuration files, this software can potentially be integrated into any robotic platform with ease and flexibility. The architecture itself is fairly straightforward, and consists of: *TAME Manager* (the main module of the system), *TAME Communication Manager* (receives sensor data and passes the updated affective values to the robot), a module for each of the affective components, and *Stimuli Interpreter*.

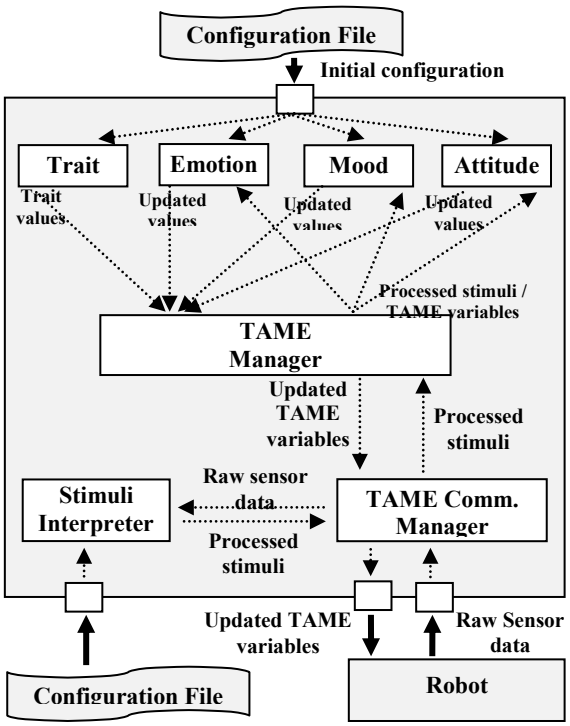

Figure 9: High-level architectural view of the TAME system.

4.1 Affective Modules

These are comprised of four different affective components of TAME (namely *Trait*, *Attitude*, *Mood*, and *Emotion*), and each module processes sensory and internal information (current values of other affective components) and calculates the updated affective variables, passing them along to *TAME Manager*. In order to provide flexibility and adaptation to individual users and situations, each component is loaded with some initial default values from a configuration file.

For the *Trait* component, a default value can be specified for each of the five personality dimensions: *Openness, Agreeableness*, *Conscientiousness*, *Extraversion*, and *Neuroticism*. Once the values are specified, they remain unchanged throughout the lifetime of the system execution since personality is generally regarded to be time-invariant.

For the *Emotion* component, configuration settings include specifications prescribing how each personality dimension may influence each emotion (e.g., directly or indirectly), as well as minimum and maximum values for a number of variables used for emotion generation. Although it would be possible for an advanced user to select these values to suit a particular task, in general, setting these defaults would be best left to the designer or administrator, as they would influence complex interactions within the module.

For the *Mood* component, configuration settings include settings for high and low points for circadian changes, as well as mapping between sensor data types and influence on mood generation. Again, some of these settings are best left for designer/administrator, except for

specifying circadian changes to match user's mood. Reinforcement learning is being investigated as the basis for determining these parameters.

Finally, for the *Attitude* component, the configuration file can specify the types and properties of objects towards which the robot can have positive or negative attitudes. Case-based learning is being developed for this component as a means to both set the parameters and generalize over broad classes of objects.

4.2 TAME Manager and TAME Communication Manager

TAME Manager is the main module in the system that runs as a threaded process to manage all the affective components. It supplies each of them with relevant sensor data (processed as stimuli) or necessary values of certain variables from other affective components. The affective variables of all the affective modules such as the *Joy* variable in the *Emotion* component or the *Extraversion* variable in the *Trait* component are comprehensively called the TAME variables in the system. Then, *TAME Manager* receives the updated values of the TAME variables for each affective component after appropriate calculations.

TAME Communication Manager is a separate thread that is responsible for receiving sensor data from the robot and relaying them to *Stimuli Interpreter*, and then passing appropriately processed stimuli into *TAME Manager*. It also receives the most up-to-date values of the TAME variables from *TAME Manager* and communicates the information to the robot controller.

Behavioral arbitration or the changes to behavioral parameters according to different affective states can then be achieved on the robot controller side. By avoiding direct manipulation of behavioral parameters, the design of our affective system allows for high portability and scalability. On the robot side, depending on the capabilities of a particular platform, corresponding affect can be implemented in either continuous or discrete manner. For example, in the continuous case, an emotion can be expressed in accordance with Ekman's FACS on an animated robot face, or mapped onto velocity and expensiveness of gestures and posture, in a manner similar to that proposed by the designers of Robovie [16] or through behavioral overlay method proposed by Brooks et al. [7]. In the discrete case, a number of affective expressions (facial and/or bodily), perhaps differing in intensity, can be designed on the robot a priory, and then an appropriate expression can be selected based on the actual value of a TAME parameter. We have implemented the latter case on a humanoid robot Nao.

4.3 Stimuli Interpreter

The raw sensor data from the robot themselves are useless unless some context is provided for them. A configuration file for *Stimuli Interpreter* gives such contextual information. For each TAME variable that is directly affected by environmental conditions or stimuli (all but traits), the file specifies whether each type of incoming sensor data is relevant to the calculation of that TAME variable. As emotions are invoked in response to specific stimuli, certain object properties would be used for stimulus strength calculation. These properties may correspond to preprocessed incoming sensor data, such as distance, size, approach angle and acceleration, or color of an object; they can also include more abstract properties, such as friendliness or disapproval of a person. A configuration file would specify which of these are relevant for generating a particular emotion, as well as weights describing a relative importance of each. For example, for fear, object size and speed of approach may play a larger role, whereas an interacting person's personal attributes may be more important in case of joy. For moods, incoming external and internal sensor data would include battery level, internal and external temperature, brightness and noise level, and other potential influences. For example, positive mood is more susceptible to energy consumption, and negative to lighting conditions, and these differences are reflected through assigning appropriate weight for each in the configuration file. Finally, for attitudes, an object identifier is used (such as an AR marker), which encode specific object properties: color, size, shape, category and material.

If a type of sensor data is relevant, a scaling factor is provided for normalization purposes, to translate it to an appropriate strength since each type can have a value in different ranges. The configuration file also allows flexibility in specifying whether multiple sensor types should have a combined effect as an average (smoothing) or the stimulus with the maximum strength should have a dominating effect in generating the value of each TAME variable (winner-take-all).

5 Implementation

The *TAME Module* was incorporated into *MissionLab*, a robotic software toolset that allows an operator to easily create and configure a multi-robot mission using a graphical user interface [36, 37][1]. In order to demonstrate the effectiveness of our affective system, it has been initially tested using Aldebaran Robotics' Nao humanoid platform (Figure 10).

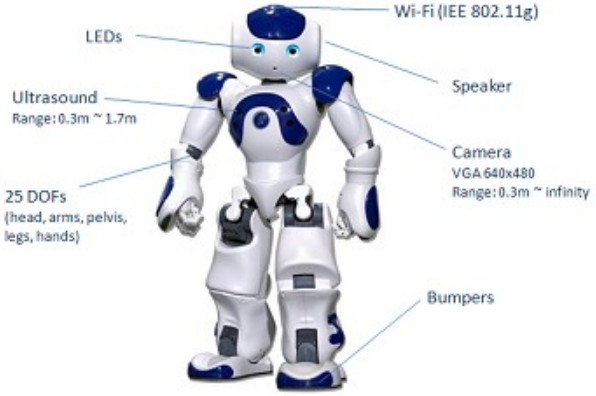

Figure 10: Aldebaran Robotics' Nao humanoid robot (source Aldebaran Robotics)

5.1 *MissionLab* Overview

In *MissionLab*, an operator uses the *Configuration Editor* (*cfgEdit*) to specify a mission using a graphical representation called an FSA, or finite state acceptor [25]. In FSA representation, a mission is composed of a combination of various actions (behaviors) to perform, and perceptual triggers act as conditions for moving from one action to the next. The resulting mission is translated into C++ code and compiled to make *Robot Executable*. Then, it can be deployed on a wide variety of simulated and real robot platforms, and the operator can monitor the execution of the mission in real-time using *mlab* GUI display. *HServer* [37] is a control interface to a variety of robotic hardware, and it is separate from *Robot Executable* to enable more flexible coordination with different robotic platforms.

5.2 Integration with MissionLab and Nao Robot

Figure 11 presents a graphical view of the integration. Here, *HServer* acts as a bridge to the Nao robot to communicate between *Robot Executable* (which contains the actual control code for the robot's current mission) and the *TAME Module*. In *HServer*, an interface for the Nao robot has been created using Nao's API for hardware control. When *Robot Executable* is in a certain behavioral state within a given mission, the generated motor commands are transmitted to *HServer*, which controls the Nao robot at the hardware level.

HServer also continuously receives perceptual data from the robot. Upon receiving the data, *HServer* sends them to both *Robot Executable* and the *TAME Module*. *Robot Executable* needs the sensor data for performing certain behaviors and for determining when to transition from one state to the next in the mission. When sending the sensor data to the *TAME Module*, *HServer* organizes relevant data for the *TAME module* in accordance with the configuration file for *Stimuli Interpreter*, and sends each type of sensor data with a unique ID.

The *TAME Module* interprets each datum in context using its *Stimuli Interpreter* and then the updated values of its TAME variables are calculated accordingly. *Robot Executable* possesses a simple database of the TAME variables, and their values are updated at 3 hertz (to ease computational burden) by the *TAME Module*. These variables influence the robot's behaviors by changing appropriate behavioral parameters or selecting from a predefined set of expressive affective behaviors.

[1] MissionLab is freely available for research and development and can be found at http://www.cc.gatech.edu/ai/robot-lab/research/MissionLab/

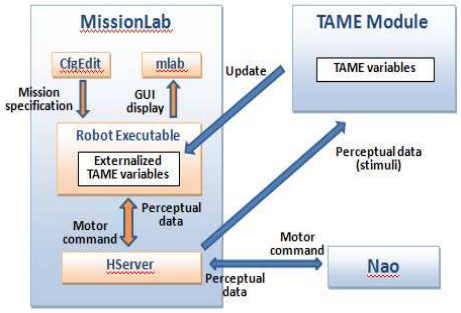

Figure 11: Architectural view of the *TAME Module* integrated with *MissionLab* and Nao humanoid robot.

5.3 Nonverbal Affective Behavior Recognition Survey

All components have been implemented on an Aldebaran Nao robot. Based on an extensive literature review [38-46], we designed expressions of Extraversion and Introversion, Positive and Negative Mood, and Emotions of Fear and Joy. Figure 12 (*Left*) and (*Right*) provides examples of static poses of Joy and Fear, respectively. To test the recognition of these affective behaviors, we conducted an online survey, in which 26 participants were asked to watch a number of short videos of Nao producing the aforementioned affective expressions [47]. The following measures were used in this survey: a shortened version of PANAS (mood) questionnaire [35] to assess Negative and Positive Mood of the robot (1 clip with the robot displaying Negative Mood and 1 - Positive Mood); Extraversion subset of Mini-Markers Big-Five personality questionnaire [34] to assess to what extent the robot in the corresponding two clips was perceived as extraverted or introverted; and a multiple-choice question asking the participants to select one of six emotions (Anger, Joy, Interest, Fear, Disgust and Sadness) or suggest another one if not present in the choice, for the clips with the robot exhibiting Joy and Fear. To the best of our knowledge, there has been no explicit research that addresses nonverbal behavior for humanoids across multiple affective constructs.

Figure 12: (*Left*) Static pose for Joy. (*Right*) Static pose for Fear.

On the Extraversion scale from 1 to 9, the Introverted Nao scored 3.6, and Extraverted – 7.1 (almost twice as Extraverted); this result was statistically significant ($p<0.001$), see Table 2 for Mean and Standard Deviation. In terms of mood expressions, the robot displaying positive mood was rated low on Negative and high on Positive Affect; the robot displaying negative mood was rated medium on Negative and low-medium on Positive Affect. For the positive robot mood, Positive Affect score was significantly higher than that for the negative robot mood (21 vs. 12.3 out of 30, $p<0.001$), and vice versa, its Negative Affect score was significantly lower than that of negative robot mood (8.6 vs. 12.3 out of 30, $p<0.001$); see Table 2 for Mean and STD. Finally, the recognition rates for emotions of joy and fear were high – 85% and 81%, respectively; these rates are comparable to those obtained in judgments of joy and fear portrayals by human actors in movie clips (facial features obscured), which were 87% and 91%, respectively [48]. Given this successful encoding of a number of affective behaviors, we are currently designing and conducting a set of human-robot interaction studies to test the effect of the system on physically present users. In the interim, a number of videos demonstrating the results to date (including the videos used for the survey) are available at: http://www.cc.gatech.edu/ai/robot-lab/tame/index.html#multi.

	Extraverted Robot	Introverted Robot	Negative Affect/ Negative Mood Clip	Positive Affect/ Negative Mood Clip	Negative Affect/ Positive Mood Clip	Positive Affect/ Positive Mood Clip
Mean	7.1	3.6	15.6	12.3	8.6	21
SD	1.1	1.2	4.5	4.1	4.1	4.8

Table 2. Mean Scores and Standard Deviations for Personality and Mood

6 Conclusion

In this paper, design and theoretical foundations have been presented for a comprehensive affective robotic architecture, spanning four different affective phenomena (namely, personality Traits, Attitudes, Moods and Emotions) that vary greatly in time, function and object specificity. As humans are highly susceptible to affective cues, and use this information, consciously or subconsciously, in everyday decisions, we believe that providing an extensive affective framework capable of generating a wealth of expressive affective behaviors would facilitate human-robot interaction, especially on a continuous and long-term basis. Some of the expected advantages of a comprehensive affective robotic system would include: increased compliance with the robot's requests, due to a better understanding and more natural interaction; greater satisfaction and bonding with the robot due to affective affinity; greater robot acceptance and reduced learning effort by making the interactions more intuitive. Our initial results with the framework on AIBO suggest that people may prefer robots with emotional capabilities, and the survey on recognition of affective robotic behavior demonstrates the plausibility of successful manifestation of affect in humanoids using TAME. As with any complex system, extensive, multi-tier studies with real people are required to thoroughly assess its potential.

Towards that end, we are currently designing and performing a series of human-robot studies to examine how each of the adaptive affective components affects everyday human-robot interaction. The first study in the series evaluates the effect of display of Negative Mood and Emotions (Fear) on subjects' compliance with a robot's request to evacuate a "dangerous" zone within a context of a mock-up search-and-rescue scenario. This study also measures the perceptions of the robot's persuasiveness and naturalness, and correlates them with an objective measure of compliance (if, and how fast the subject leaves a danger zone). The second study is designed to examine the value of robotic personality traits (Extraversion vs. Introversion) for different types of tasks in the context of a building demolition exhibit. Additionally, a longitudinal study is being composed to assess the effectiveness of more subtle phenomena, such as attitudes and moods on long-term human-robot interaction in an office setting. As part of these studies, we are also planning to develop a wider repertoire of affective displays for humanoid robots, and conduct an extended online or real-robot survey to determine the recognition rates. Finally, we intend to use the framework as a test-bed for new objective and subjective performance measures of affective robotic systems.

References

[1] Moshkina, L., Arkin R.C., Lee, J.K. and Jung, H., *Time-varying affective response for humanoid robots*, IEEE International Conference on Social Robotics, 2009.

[2] Moshkina, L., *An Integrative Framework for Affective Agent Behavior*, IASTED International Conference on Intelligent Systems and Control, 2006.

[3] Moshkina, L., Arkin, R.C. *Human Perspective on Affective Robotic Behavior: A Longitudinal Study*, IEEE IROS, 2005.

[4] Carney, D.R., Colvin, C.R., and Hall, J.A., *A thin slice perspective on the accuracy of first impressions,* Journal of Research in Personality 41, pp 1054-1072, 2007.

[5] Nass, C., Moon, Y., Morkes, J., Kim, E-Y., Fogg, B.J., *Computers are social actors: A review of current research*, in *Moral and ethical issues in human-computer interaction*, B. Friedman, Editor. CSLI Press: Stanford, CA,.1997.

[6] Arkin, R.C., *Moving Up the Food Chain: Motivation and Emotion in Behavior-based Robots*, in Who Needs Emotions: The Brain Meets the Robot (Fellous, J. and Arbib, M., Editors), Oxford University Press, 2005.

[7] Brooks, A. and Arkin, R.C., *Behavioral Overlays for Non-Verbal Communication Expression on a Humanoid Robot*, Autonomous Robots, Vol. 22, No.1, pp. 55-75, Jan. 2007.

[8] Arkin, R., Fujita, M., Takagi, T., and Hasegawa, R., *An Ethological and Emotional Basis for Human-Robot Interaction*, Robotics and Autonomous Systems, 42 (3-4), March 2003.

[9] Miwa, H., Takanishi, A., Takanobu, H., *Experimental Study on Robot Personality for Humanoid Head Robot,* IEEE IROS, p. 1183-1188, 2001.
[10] Gockley, R., Simmons, R., and Forlizzi, J., *Modeling Affect in Socially Interactive Robots*, IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06), Hatfield, UK, September 6-8, 2006
[11] Fukuda, T., Jung, M., Nakashima, M., Arai, F., Hasegawa, Y., *Facial expressive robotic head system for human-robot communication and its application in home environment*, Proceedings of the IEEE, p. 1851-1865, 2004.
[12] Breazeal, C., *Emotion and Sociable Humanoid Robots,* International Journal of Human-Computer Studies, p. 119-155, 2003.
[13] Esau, N., Kleinjohann, B., Kleinjohann, L., Stichling, D., *MEXI: machine with emotionally eXtended intelligence,* Design and application of hybrid intelligent systems, p. 961-970, 2003.
[14] Murray, J.C., Canamero, L. and Hiolle, A., *Towards a Model of Emotion Expression in an Interactive Robot Head*, IEEE International Symposium on Robot and Human Interactive Communication, 2009.
[15] Canamero, L., *Playing the Emotion Game with Felix: What Can a Lego Robot Tell Us about Emotion?* In *Socially Intelligent Agents*, K. Dautenhahn, Bond, A.H., Canamero, L., Edmonds, B., Editor. Kluwer Academic Publishers, 2002.
[16] Nakagawa, K.; Shinozawa, K.; Ishiguro, H.; Akimoto, T.; Hagita, N., *Motion modification method to control affective nuances for robots*, IEEE IROS, 2009
[17] Wu, T., Butko, N.J., Ruvulo, P., Bartlett, M.S., and Movellan, J.R., *Learning to Make Facial Expressions*, IEEE International Conference on Development and Learning, 2009
[18] P. Ekman and W. Friesen, *Facial Action Coding System (FACS): A technique for the measurement of facial action*, Palo Alto, CA: Consulting, 1978.
[19] Goris, K., Saldien, J., and Lefeber D., *Probo, a testbed for Human Robot Interaction*, ACM/IEEE International Conference on Human-Robot Interaction, 2009.
[20] Sosnowski, S., Bittermann, A., Kuhnlenz, K., and Buss, M., *Design and Evaluation of Emotion-Display EDDIE*, IEEE IROS, 2006.
[21] Lee, K.M., Peng, W., Jin, S.A., and Yan, C., *Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human-robot interaction,* Journal of Communication, 56 (4), pp 754-772, 2006.
[22] Clore, G., Clore, G. L., & Palmer, J., *Affective guidance of intelligent agents: How emotion controls cognition*, Cognitive Systems Research*,10,* 21–30, 2009.
[23] Moffat, D., *Personality Parameters and Programs,* Creating Personalities for Synthetic Actors, p. 120-165, 1997.
[24] Watson, D., Clark, L.A., Tellegen, A., *Mood and Temperament*, The Guilford Press, 2000.
[25] Arkin, R.C., *Behavior-based Robotics*, MIT Press, 1998.
[26] McCrae, R.R., Costa, P.T., *Toward a new generation of personality theories: theoretical contexts for the five-factor model*, Five-Factor Model of Personality, p. 51-87, 1996.
[27] Costa, P.T., McCrae, R.R., *NEO PI-R Professional Manual,* Psychological Assessment Resources, 1992.
[28] Ekman, P., *Moods, Emotions and Traits*, in *Nature of Emotion: Fundamental Questions*, Ekman P., R., Editor. 1994, Oxford University Press: New York.
[29] Picard, R.W., *Affective Computing*, MIT Press, 1997.
[30] Davidson, R.J., *Affective Styles and Their Cognitive Consequences*, Handbook of Cognition and Emotion (T. Dalgleish, Power, M.J., Editors), John Wiley and Sons, p. 103-124, 1999.
[31] Frijda, N.H., *Varieties of Affect: Emotions and Episodes, Moods, and Sentiments*, in The Nature of Emotion: Fundamental Questions (P. Ekman, Davidson, R.J., Editors), Oxford University Press, 1994.
[32] Breckler, S.J., Wiggins, E.C., *On Defining Attitude and Attitude Theory: Once More with Feeling*, Attitude Structure and Function (A.R. Pratkanis, Breckler, S.J., Greenwald, A.G., Editors), Lawrence Erlbaum Associates, p. 407-429, 1989.
[33] A. Aamodt, E. Plaza, *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches*. AI Communications, IOS Press, Vol. 7: 1, pp. 39-59, 1994.
[34] Saucier, G., *Mini-Markers: a Brief Version of Goldberg's Unipolar Big-Five Markers.* Journal of Personality Assessment, 1994. **63**(3): p. 506-516.
[35] Watson, D., Clark, L.A., Tellegen, A., *Development and validation of brief measures of positive and negative affect: The PANAS scales.* Journal of Personality & Social Psychology, 1998. **56**(6): p. 1063-1070.
[36] D. MacKenzie, R. Arkin, and J. Cameron, *Multiagent mission specification and execution*, Autonomous Robots, vol. 4, no. 1, pp. 29-52, March 1997.
[37] Georgia Tech Mobile Robot Laboratory, Manual for MissionLab Version 7.0, 2007. Available: http://www.cc.gatech.edu/ai/robot-lab/research/MissionLab/
[38] Nass, C. and Lee, K.M., *Does computer-generated speech manifest personality? An Experimental Test of Similarity-Attraction*, ACM CHI 2000.
[39] Lee, K.M., Peng, W., Jin, S.A., and Yan, C., *Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human-robot interaction,* Journal of Communication, 56 (4), pp 754-772, 2006.
[40] Gallaher, P. E., *Individual differences in nonverbal behavior: Dimensions of style*, Journal of Personality and Social Psychology, 63, 133–145, 1992.
[41] Gifford, R., *Mapping nonverbal behavior on the interpersonal circle*. Journal of Personality & Social Psychology, 61, 279–288, 1991.

[42] Atkinson, A.P., Dittrich, W.H., Gemmell, A.J., Young, A.W., *Emotion perception from dynamic and static body expressions in point-light and full-light displays*. Perception 33 (6), 717–746, 2004.
[43] Wallbott, H.G.: *Bodily expression of emotion*. European Journal of Social Psychology 28 879–896, 1998.
[44] Coulson, M., *Expressing emotion through body movement: A component process approach*, in: R. Aylett, L. Cañamero (Eds.), Animating Expressive Characters for Social Interactions, SSAISB Press, 2002.
[45] Mehrabian, A., Ksionsky, R., *A theory of Affiliation*. Lexington, MA: Heath, 1974
[46] Argyle, M., *Bodily Communication*. London, U.K.: Methuen, 1975
[47] Park, S., Moshkina, L., and Arkin, R.C., *Recognizing Nonverbal Affective Behavior in Humanoid Robots*, International Conference on Intelligent Autonomous Systems, Ottawa, Canada, 2010
[48] Atkinson, A.P., Dittrich, W.H., Gemmell, A.J., Young, A.W., *Emotion perception from dynamic and static body expressions in point-light and full-light displays*. Perception 33 (6), 717–746, 2004.