Situation-Based Stimuli Response in a Humanoid Robot

Palis Ratanaswasd, Chris Garber, and Adrian Lauf

Electrical Engineering and Computer Science
Vanderbilt University
Nashville, TN 37235-0131, USA

{palis.ratanaswasd, chris.garber, adrian.p.lauf}@vanderbilt.edu

Abstract **–Implementation of human cognitive abilities has become a key area of interest to robotic researchers as cognitive abilities are believed to be linked to high-level intelligent behaviors. In this paper, we associated emotions with stimuli to help the humanoid robot ISAC decide his action when responding to a particular situation. According to cognitive psychology, sound stimuli can create different kinds of emotional response in humans. Therefore, we have designed ISAC to possess different kinds of emotional response according to different sounds being perceived. Emotion, along with attention, are used by the decision making process to decide actions to be performed. Decisions are made based on the knowledge learned about the situations. The simulations in this paper show how these human cognitive abilities can be used to generate different responses to sound stimuli depending on the current situation.**

Index Terms - humanoid robot, cognitive control, attention, emotion, situation-based

I. Introduction

Traditionally, robot planning systems are implemented using deterministic methods to decide which action to execute based on sensory information. This deliberative approach works well in closed, well structured environments, such as in a research lab. However, humanoid robots that interact with humans in dynamic environments need more robust decision-making capabilities to allow different actions to be executed given the same set of sensory information in different situations. Decisions to perform specific actions must be based on the situation of the robot at the current moment.

In contrast to many robots, humans and most animals have the ability to make suitable decisions in different situations. Often, we find emotion is one of the factors. Emotion not only affects the decision but also affects how an action will be performed, e.g., a person tends to do things more aggressively when they are angry.

Emotion can be used as the bias for decisions during a robot's decision-making process. This is illustrated using an example involving a dog shown in Fig. 1. In this situation, the dog needs to decide whether he wants to approach the bone, escape, attack the human, or ignore everything. The dog uses his emotional response to his percepts to decide what action to take. If the dog fears the human enough, he will suppress his hunger and ignore the bone.

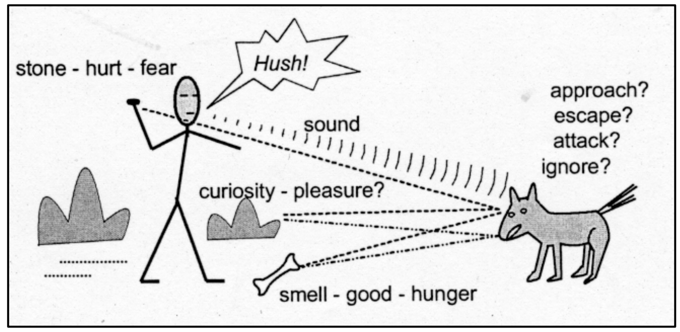

Fig. 1. An example of how emotions play a role in action decision based on situation [1]

Recently, robotic researchers have become more interested in human cognitive abilities such as attention [2][3] and emotion [4][5]. Knowledge in cognitive science is applied to robots in different ways to enhance their capability in task execution. One effort in our research is to develop a cognitive robot equipped with human inspired cognitive control, emotion, and attention, to allow it to be fluent in both reactive responses and deliberative actions.

II. Human Cognitive Abilities

A. *Information processing/ response to stimuli*

The general concept of human cognition involves the processing of information within the brain. This processing begins when a human perceives stimuli. In this perception stage, information is encoded into representations or percepts. The encoded sensory information is retained briefly in this stage before being passed into the attention stage. At the attention stage, percepts are sent through a gating mechanism to be passed into the working memory (also known by some cognitive science researchers as *short-term memory* [6][7]). Percepts that are allowed to pass through the gating mechanism are said to have received "spatial attention" [8]. Information from this stage will be further used by cognitive processes during task execution.

B. *Attention*

Attention is a cognitive process that limits the amount of information from perception systems to be manipulated by the brain. It allows the brain to concentrate only on particular information by filtering out distracters from a desired target object or spatial location by amplification of the target representations [9]. Attention can be selective during task executions, such as when searching for an object, and, at the same time, can be automatic in salience events such as when hearing a loud sound. Caused by characteristics of stimuli, an arousal can raise or lower the overall level of attention, thereby allowing attention to be focused on a certain task or environmental stimuli [10]. According to Neisser [11], attention is influenced by two factors: properties of the stimuli and existing knowledge, which the latter includes past experience. In addition, attention and emotion can influence each other and can be caused by the same percepts [9]. For example, a loud sound can shock a person, creating surprise, and therefore cause him to pay attention to it.

C. *Emotion*

Emotions play a critical role as percept value judgments during decision making about task switching in a variety of situations. Emotional reactions to percepts are value state-variables that are used to evaluate the benefits and costs of plans and actions based on the state of the world [12]. One example is that humans have the ability to shift focus of attention from the current task being performed based on the automatic processing of unattended acoustic stimuli. For instance, hearing one's name spoken from across a crowded room where many conversations are taking place while listening to another conversation simultaneously.

While intensity of acoustic stimuli contributes modestly to physical response, emotional reaction to acoustic stimuli contributes strongly to physical response [13]. If the physical response is strong enough it could cause a shift in focus of attention during task execution. For example, continuous sinusoidal tones, like alarm signals, sound irritating and cause displeasure. Limited duration sinusoidal tones that contain harmonics, like signals from musical instruments, sound pleasant and cause pleasure [1]. Thus, alarm signals cause strong emotional reaction leading to a shift of attention. An example of two alarm sounds that are irritating to human ears causing displeasure is shown in Fig. 2. These graphs illustrate pulses at certain frequencies indicating pure sine tones as main features of the sounds.

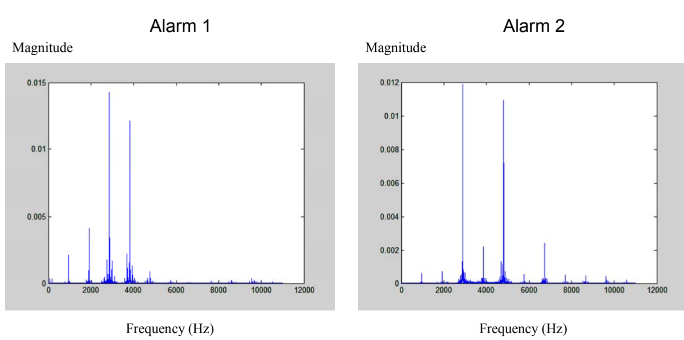

Fig. 2. Frequency responses of sounds from two alarm clocks

The link between emotional reaction to acoustic stimuli and focus of attention during task execution is a key concept when attempting to implement cognition in a sensorimotor intelligent system. Attention and emotion are critical components of intelligence that are necessary for sensory input evaluation, goal selection, directing behavior, and controlling learning [12].

D. *Cognitive control*

The human brain contains a vast amount of information such as perceptual, semantic, motoric, etc. During task execution only a small set of these resources are needed. Human cognitive skills involve the ability to organize resources necessary for the task at hand, including selection and maintenance of required information and avoidance of disruption from other influences [14]. Disruptions can come in the form of sensorimotor type actions which tend to be triggered in response to certain stimuli regardless of current goals. In order to accomplish particular goals, execution of actions must be controlled according to the context. Actions are selected based on information obtained from past experiences and knowledge about the task and environment. In some cases, sensorimotor coordination must be either inhibited or overridden. The ability to control actions in this manner to accomplish a task is called cognitive control [7][14][15][16][17]. Effective cognitive control requires a subtle manipulation of task switching that prevents disruption of ongoing tasks while providing flexibility to allow rapid execution of other tasks that arise in a given situation [18]. The concept of cognitive control is illustrated in Fig. 3.

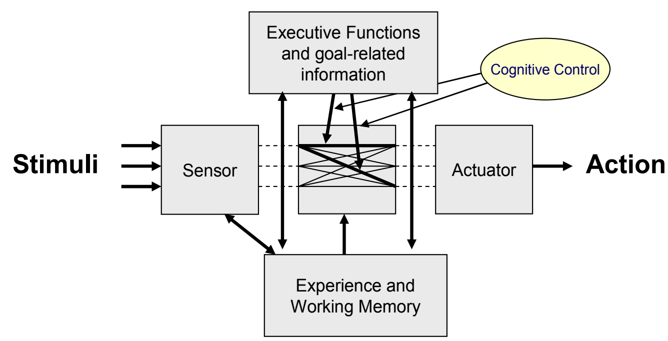

Fig. 3 Concept of cognitive control, based on [16]

According to Miller, executive functions of cognitive control include the ability of the brain to [16]:

- Generate plans and monitor task progress
- Focus on task related information
- Maintain and update goal information
- Inhibit distractions to current tasks
- Shift between different levels of cognition ranging from routine actions to complex deliberation
- Learn new responses in novel situations.

III. Architecture

Implementation of human cognitive control, attention, and emotion is being applied on the humanoid robot ISAC (Intelligent Soft Arm Control) developed at the Center for Intelligent Systems at Vanderbilt University. ISAC is a general-purpose humanoid robot with two 6-DOF arms, hands, various sensors, and an active stereo vision system (with pan-tilt capability) mounted on the head. Sound is sensed through a pair of microphones which also allow for the detection of the direction of the sound source. Other sensors mounted on ISAC include infrared and touch sensors.

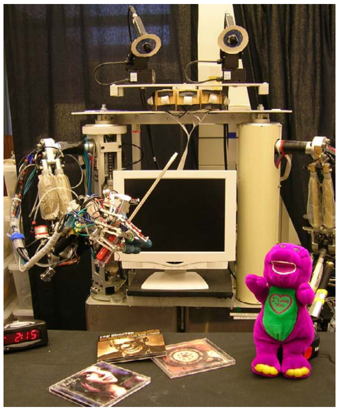

Fig. 4. ISAC Humanoid Robot

The software architecture used on ISAC is a multiagent-based system that incorporates several components inspired by different cognitive processes in humans called the Intelligent Machine Architecture (IMA) [19][20]. ISAC's cognitive abilities are implemented as collections of IMA agents and memory structures. This architecture is shown in Fig. 5.

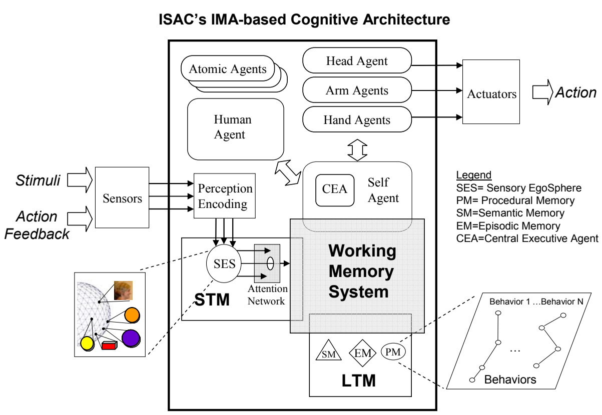

Fig. 5. ISAC's Multiagent-based Cognitive Robot Architecture

Memory structures are utilized in ISAC's cognitive robot architecture to help maintain the information necessary for immediate tasks and to store experiences that can be used during decision making processes.

Short-term memory (STM) stores sensory information on a structure called the Sensory Ego Sphere (SES) [21]. Memories in SES can be retrieved using stimulus content such as key words, colors, or time of posting. The stored information decays over time.

Long-term memory (LTM) stores information such as learned skills, semantic knowledge, and past experience for retrieval in the future. As a part of LTM, Procedural Memory (PM) holds motion primitives and behaviors needed for movement, such as how to *reach to a point* [22]. Behaviors are derived using the Spatio-Temporal Isomap method proposed by Jenkins and Matarić [23]. Semantic Memory (SM) is a data structure composed of known objects in the environment. Episodic Memory (EM) stores past experiences including goals and task sequences that ISAC has performed in the past. EM implementation is described in detail by Dodd and Gutierrez in [24].

The Working Memory System (WMS) is modeled after the working memory in humans, which holds a limited number of "chunks" of information needed to perform a task, such as a phone number during a phone dialing task. It allows the robot to focus attention on the most relevant features of the current task, which is closely tied to the learning and execution of tasks [25]. In our implementation, the WMS consists of two functional parts: a set of working memory slots that can be accessed by other agents, and a device to select chunks of memory to place in those slots based on the current task context. WMS works with the Central Executive Agent (CEA) to provide cognitive control similar to that seen in humans. This working memory model is based on models from computational neuroscience which are discussed in detail by Skubic et al. in [26]. Please see [27] for more discussion on WMS.

IV. APPROACH

In this section, we describe the method used by ISAC to make decisions when encountering conflicting goals during a specific situation. A variety of information is used during the decision making process, including the context of the current task, environmental conditions, and past experiences.

A. *Attention*

Represented in the form of percepts, sensory information of the immediate environment is essential for the robot to perform tasks properly. Although most percepts can be ignored, some should receive attention by the system, including those required for task execution and those with high emotional salience.

ISAC pays attention to the percepts in the environment through the Attention Network [21]. Recognized percepts are posted on SES. The Attention Network focuses attention on the percepts by making the directions of these percepts known to other processes. This procedure is known as providing *focus of attention* (FOA). The percepts that receive FOA are considered a part of WMS and will be used by CEA during task execution as shown in Fig. 6.

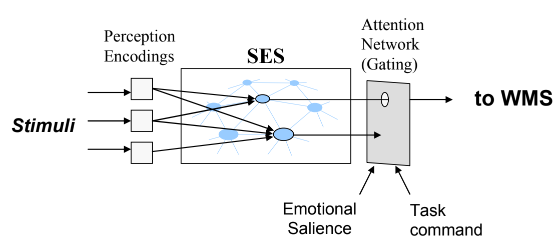

Fig 6. Focus of Attention (FOA)

B. Emotion

In ISAC's cognitive architecture, emotions are managed by the Emotion Agent (EA). Dodd and Gutierrez described the initial design of the EA in [24]. There are four levels in this design, with the lowest level being Elementary Sensation. Sensations can be perceived from outside through sensors as well as from the internal state of the robot. These elementary sensations are then mapped to a basis of response using a network called the Emotion System. The output of the Emotion System is a vector representing the emotional state of ISAC. Each element of the vector has a magnitude that varies between -1 to 1. The magnitude of this vector is sent to the Attention Network as the level of emotional salience for the given stimulus or task.

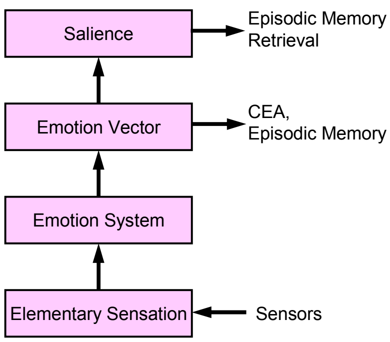

Fig.7. Overview of Emotion Agent [24]

In the current implementation, the Emotion System is highly active to a number of sound sensations. These responses are implemented as initial knowledge based on findings of human emotional response of various sounds described in section II.B. Percepts that contain sensations that cause high emotional salience from the Emotion Agent will receive FOA from the Attention Network.

Stimuli are processed with various signal processing techniques in order to find their corresponding elementary sensations. For example, Fourier analysis can be used to detect frequencies that create irritating feelings in alarm sounds. We have analyzed a few samples of sound percepts and mapped them into Emotion Vectors. The following table shows an example of emotional responses to various types of stimuli.

TABLE 1
STIMULI/TASKS VS. CORRESPONDING EMOTION VECTORS

Stimulus/Task-response	Emotion Vector			Emotional salience
	happy-sad	love-hate	alert-uninterested	
Alarm sound	0.1	-0.8	-0.6	1.01
Musical piece	0.7	0.5	0.1	0.86
Task command	0.1	0	0.6	0.61
Other human words	0.1	0.1	0.1	0.17

C. Cognitive Control

ISAC's cognitive control is modeled and implemented based on Baddeley's psychological human working memory model. His working memory model consists of the "central executive," which controls two working memory systems namely the phonological loop and the visuo-spatial sketchpad. Cognitive control in ISAC is implemented using CEA which interfaces with the WMS. Functions of CEA include the decision making of which action to execute and the monitoring of the state of the system.

To respond to stimuli, the Attention Network assigns FOAs to percepts on the SES based on their corresponding emotion vector. As CEA decides which FOA to respond to based on the current situation, a proper response to the selected FOA will be invoked. Among candidate behaviors stored in LTM as initial knowledge, a behavior is selected in response to a FOA according to stimuli-response mapping constructed and maintained by CEA. After the action is executed, feedback from the execution will be used by CEA to adjust the mapping of the current set of stimuli-response. Overview of cognitive control in ISAC is illustrated in Fig. 8.

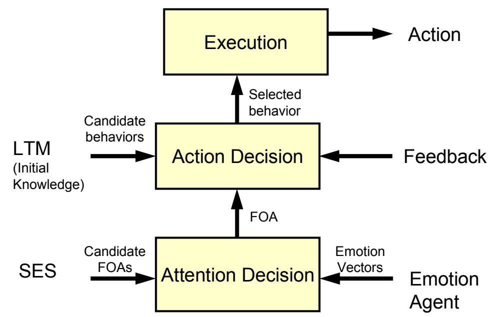

Fig. 8. Overview of cognitive control in ISAC

D. Decision Making

Conflicting goals happen when two or more percepts and/or tasks are given to CEA at the same time. For example, a percept with high emotional salience is detected while CEA is executing a task. When this happens, CEA must make a decision of how the newly acquired percept should be responded to.

In the first step, the Attention Network selects percepts to assign FOAs. This decision is made based on the level of emotional salience of a particular percept. The one with the largest emotional salience level will be selected for response. Next, CEA will respond to the percept based on initial knowledge in LTM and information learned about that particular percept. For example, CEA might know how to move an object by pushing it but later learns that the object could be moved by picking it up also.

The decision making process is constructed using a Markov model shown in Fig. 9. In this model, "situations' are considered as a series of *states* that the model has been presented. To respond to a stimulus, a *state* is represented by a set of stimuli requiring attention at a particular moment. The model changes its state when a new percept requests attention. The decision is made using the probability that an action will be chosen to handle the current situation. CEA requires a method to evaluate this decision. In this paper, human feedback is used to evaluate a task execution. In the future, we will incorporate semantic knowledge that will contain expected task execution results. The comparison between an expected and the actual results will be used as feedback.

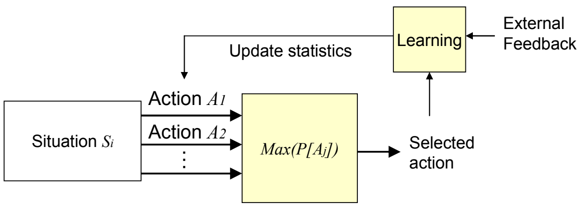

Fig. 9. Decision making and learning

V. Experiments

We have designed and ran a simulation to test the key concepts proposed in this paper for the implementation of cognitive control in a system that encompasses CEA, attention, and emotion. The simulation was set up to evaluate the ability of the system to focus attention towards percepts that cause high emotional salience, respond to percepts differently based on the current situation, and switch between tasks based on the decision made by the CEA.

The simulation uses a set of initial knowledge about acoustic stimuli and actions. Sound stimuli were captured through a microphone and processed in Matlab. The model could recognize a few sounds including, "music" and "alarm" signals. A specific piece of music was used for the "music," and a recorded sound of the actual Vanderbilt Engineering building fire alarm was used for the "alarm." Task commands were recognized using simple keyword matching, e.g. "reach to [an object]" or "pick up [an object]". The model was set up to dislike the alarm sound while loving the presented piece of music. Emotional response to stimuli and tasks represented by emotion vectors are shown in Table I. The values in the emotion vectors were obtained by averaging the result from a small survey which asked each subject to rate each emotional response ranging from -1 to 1. CEA, Attention Network, and Emotion Agent were also modeled using Matlab. The model was assumed to be capable of performing three actions, performing the requested task, yelling "Alarm!", and performing a free-style robot dance. Initially, the model used these actions to respond to task commands, alarm sound, and music sound respectively.

The simulation was run using two similar situations to evaluate the decision making process. In both situations, the two sound percepts were given to the model to demonstrate its capability to focus attention. In the second situation, a task was first given to the model. Later one of the percepts was given to introduce conflicting input. The model would decide whether or not to shift attention from the task being performed. The human, who issued the task, made suggestions to the model based on the decision it previously made. The model then learned from the feedback and adjusted its decision making mechanism for similar situations in the future. The overview of this experiment is illustrated in Fig. 10.

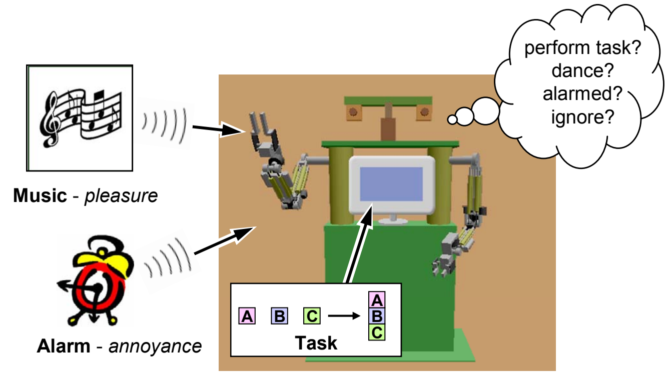

Fig. 10. Overview of the experiment

VI. Performance Evaluation

In this simulation, we evaluated the following aspects of the proposed cognitive control system:

1) The ability to focus attention towards percepts that cause high emotional salience
2) The ability to respond to percepts differently based on the current situation
3) The ability to switch between tasks based on the decision made by the CEA
4) The ability to learn to make different decisions based on the feedback from previous tasks

In the first situation, various sounds were presented to the simulation at different times including a short piece of music, an alarm sound, human voices, and background noises. Among these, the musical piece and alarm sound, and some human words were recognized. However, only the musical piece and alarm sound caused the Emotion Agent to create emotion vectors with high magnitude. The Attention Network assigned FOA by posting pointers to these two percepts on WMS making them available to CEA. In this situation, CEA used previously learned actions to respond to the sounds. Because no task was given, CEA responded using the action already learned for each percept. Therefore, this situation demonstrated the first aspect of the performance evaluation.

The second situation was similar to the first one, except a task was assigned to the model prior to the presence of other stimuli. The situations were determined to be different when the two stimuli were perceived. In the case of music, the situation was "Music was heard during a task" while it was "Alarm was heard during a task" for the case of the alarm sound. The high emotional salience of the two sounds still caused FOA. CEA decided if it should pay attention to the stimuli or keep focusing on the task based on prior knowledge of the stimuli and situation.

In situation 2, human feedback was also given to CEA so it learned whether the responses were suitable. During this trial, the human suggested that the model ignore the alarm after it previously responded by the "yelling" action. After four trials, in average, the model learned to ignore the alarm. To verify this learning process, both stimuli were presented to the model for the second time. Table II summarizes the responses of FOA in this experiment.

TABLE II
STIMULI RESPONSES

FOA	Situation 1	Situation 2 (before learning)	Situation 2 (after learning)
Music	"Dancing"	Ignored the music	Ignored the music
Alarm	Yelled "Alarm!"	Yelled "Alarm!"	Ignored the alarm

A variation of situation 2 was also performed by presenting a stimulus while the model was responding to another. By using what was learned from situation 2, the model performed "dancing" when no task was presented, and always ignored the alarm. This shows that the model learned that it should ignore the alarm altogether as a result of what was previously learned. Situation 2 demonstrated all aspects of the performance evaluation.

One problem that we foresee is that the model will take more time to learn to select other actions as it gains more experience with any one action. This is due to the fact that the probability decision of an action always becomes larger as it is selected, so it will take more number of times for other actions to be selected in order to learn a new decision. To overcome this problem, a different type of decision-making method could be used, such as Fuzzy logic or neural nets.

VII. Conclusion

We have presented a model of how to incorporate human-inspired cognitive abilities within the humanoid robot ISAC's cognitive architecture. These abilities, including attention, emotion, and cognitive control, allow the robot to behave differently in different situations. This paper focuses on using emotion to help ISAC decide actions to execute in response to certain percepts associated with emotional response. The emotions are used with the current state of the robot to decide the appropriate actions to execute. We believe the use of human inspired abilities will lead to more intelligent behaviors during robot task execution.

Acknowledgment

We would like to thank our professors, Drs. Kazuhiko Kawamura, and Mitch Wilkes for inspiration and knowledge. We also would like to thank our colleagues, Stephen Gordon, Joe Hall, and Flo Fottrell for their assistance. This work is supported in part under NSF grant EIA0325641, *"ITR: A Biologically Inspired Adaptive Working Memory System for Efficient Robot Control and Learning"*.

References

[1] Haikonen, P. O., *The Cognitive Approach to Conscious Machines*, Charlottesville, VA: Imprint Academic Press, 2003.

[2] Westin, C.F., C.J. Westelius, H. Knutsson, and G. Granlund. "Attention control for robot", *IEEE Computer Vision and Pattern Recognition*, San Francisco, CA, 18-20, 1996.

[3] Rybak, I. A., V. I. Gusakova, A.V. Golovan, L.N. Podladchikova, and N. A. Shevtsova. "A model of attention-guided visual perception and recognition," *Vision Research*, 38, 2387-2400, 1998.

[4] Breazeal, C., Emotion and sociable humanoid robots, *International Journal of Human-Computer Studies, 59*, 119-155, 2003.

[5] Gadanho, S.C., J. Hallam, "Emotion-Triggered Learning in Autonomous Robot Control," *Int'l Jo. of Cybernetics & Systems, 32*, 531-559, 2001.

[6] Ashcraft, M.H., *Cognition*, 3rd ed., Prentice Hall, 2002.

[7] Gazzaniga, M.S., R.B. Ivry, and G.R. Mangun, *Cognitive Neuroscience: The biology of the mind.* NY: Norton, 2002.

[8] Cohen, A, and R. Shoup, "Perceptual dimensional constraints on response selection processes," *Cognitive Psychology, 32,* 128-181, 1997.

[9] Taylor J.G., and N. Fragopanagos, "Modelling Human Attention and Emotions," International Joint Conference on Neural Networks, Budapest, Hungary, July 2004.

[10] Eysenck, M., *Attention and Arousal*. NY: Springer-Verlag, 1982.

[11] Neisser, U., *Cognitive Psychology*, New York: Appleton-Century-Croft, 1967.

[12] Albus, J. S., "Outline for a Theory of Inelligence," *Proc. of IEEE Systems, Man and Cybernetics,* 21(3), 473-509, May/June 1991.

[13] Bradley, M.M. and P.J. Lang, "Affective reactions to acoustic stimuli," Psychophysiology, 37, Cambridge Univ. Press, 204-215, 2000.

[14] Monsell S. and J. Driver (eds), *Control of Cognitive Processes: Attention and performance XVIII*, Cambridge, MA: MIT Press, 2000.

[15] Baars, B.J., *In the Theatre of Consciousness: The workspace of mind*, Oxford: Oxford University Press, 1996.

[16] Miller, E.K., *Cognitive Control, Fundamental of Brain and Mind Lecture Series*, Mass. Institute of Technology, June 11-13 2003.

[17] Hommel, B., K.R. Ridderinkhof, and J. Theeuwes, "Cognitive control of attention and action: Issues and trends," *Psychological Research*, 66, 215-219, 2002.

[18] Monsell, S., "Task switching," *Trends in Cognitive Sciences*, 7(3), 134-140, March 2003.

[19] Kawamura, K., R.A. Peters II, R. Bodenheimer, N. Sarkar, J. Park, and A. Spratley, "Multiagent-based Cognitive Robot Architecture and its Realization," *Int'l Jo. of Humanoid Robotics, 1(1),* 65-93, March 2004.

[20] Pack, R.T., D.M. Wilkes, and K. Kawamura, "A software architecture for integrated service robot development," *Proc. of IEEE Systems, Man and Cybernetics*, Oct. 1997, 3774-3779, 1997.

[21] Hambuchen, K.A., *Multi-Modal Attention and Binding using a Sensory EgoSphere*, Ph.D. Dissertation, Nashville, TN: Vanderbilt University, May 2004.

[22] Erol, D., J. Park, E. Turkay, K. Kawamura, O.C. Jenkins and M.J. Matarić, "Motion generation for humanoid robots with automatically derived behaviors," *Proc. of IEEE Int'l. Conf. on Systems, Man, and Cybernetics,* Washington, DC, Oct. 2003, 1816-1821, 2003.

[23] Jenkins, O.C. and M.J. Matarić, "Automated derivation of behavior vocabularies for autonomous humanoid motion," *2nd Int'l Joint Conference on Autonomous Agents and Multiagent Systems*, 2003.

[24] Dodd, W. and R. Gutierrez, "Episodic memory and emotion for a cognitive robot**,"** *14th IEEE Int'l Workshop on Robot and Human Interactive Communication (RO-MAN)*, Nashville, TN, Aug 13-15, 2005.

[25] O'Reilly, R., T.S. Braver, and J.D. Cohen, "A biologically based computational model of working memory," *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control,* (A. Miyake and P. Shah, Eds.) Cambridge: Cambridge University Press, 1999.

[26] Skubic, M., D. Noelle, M. Wilkes, K. Kawamura, and J. Keller, "A biologically inspired adaptive working memory for robots," *AAAI Fall Symp., Workshop on the Intersection of Cognitive Science and Robotics: From interfaces to intelligence*, Washington, DC, Oct. 2004.

[27] Gordon, S. and J. Hall, "System Integration with Working Memory Management for Robotic Behavior Learning", submitted to *5th International Conference on Development and Learning*, Bloomington, IN, May 31- June 3, 2006.