ISORA 2006 11th International Symposium on Robotics and Applications (ISORA) 2006
Budapest, Hungary July24-27, 2006

From Intelligent Control to Cognitive Control

Kazuhiko Kawamura and Stephen Gordon
Center for Intelligent Systems
Vanderbilt University
Nashville TN 37235-0131 USA
{kaz.kawamura; stephen.m.gordon}@vanderbilt.edu

ABSTRACT

This paper describes our efforts to develop a robot with *robust sensorimotor intelligence* using a multiagent-based robot control architecture and a biologically inspired intelligent control. Such control is called *cognitive control*. In this paper we will discuss the application of cognitive control to a humanoid robot. Features of cognitive control addressed include short-term memory for environmental learning, long-term memory for behavior learning and task execution using working memory and TD learning.

KEYWORDS: intelligent control, cognitive control, behavior learning, working memory

1. INTRODUCTION

IEEE Control Systems Society's Technical Committee on Intelligent Control states that "The area of intelligent control is a fusion of a number of research areas in systems and control, computer science, and operation research among others, coming together, merging and expanding in new directions …" [1]. A key sentence here is *fusion of research areas.* As the need to control complex systems increases, it is important to look beyond engineering and computer science areas to see if we can learn from them. For example, humans have the capacity to receive and process enormous amount of sensory information from the environment, exhibiting integrated sensorimotor intelligence as early as two years old [2]. A good example of such sensorimotor intelligence by adults is the well-known Stroop test [3, Appendix 9.1]. Thus it is a challenge for control engineers to find ways to realize human's robust sensorimotor mechanisms called *cognitive control* [4] within machines.

Most goal-oriented robots currently perform only those or similar tasks they were programmed for and very little emerging behaviors are exhibited. What is needed is an alternative paradigm for behavior learning and task execution. Specifically, we see cognitive flexibility and adaptability in the brain as desirable design goals for the next generation of intelligent robots. Several cognitive architectures have been implemented for the purpose of testing human psychological models [5][6], but such models have not been fully adopted by the robotic community.

2. INFORMATION PROCESSING IN HUMANS

Engineers have long used control systems utilizing feedback loops to control mechanical systems. Figure 1 illustrates a class of adaptive (or learning) control systems [7]. Limitations of model-based control led to a generation of *intelligent control techniques* such as fuzzy control, neuro computing and reconfigurable control [1].

The human brain is known to process a variety of stimuli in parallel, ignore non-critical stimuli to execute the task in hand, and learn new tasks with minimum assistance. This process, known as *executive* or *cognitive control*, is unique to humans and a handful of animals [2]. Figure 2 illustrates a conceptual model of cognitive control in which we are using to realize robust behavior generation and learning for our humanoid robot.

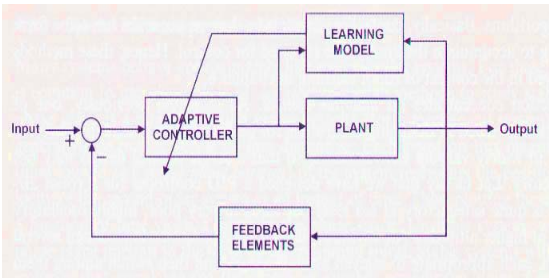


Figure 1. An adaptive control system [8]

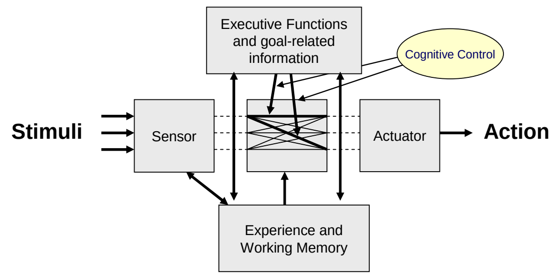


Figure 2. Model of cognitive control modified from Miller, et. al [9]

3. MULTIAGENT-BASED COGNITIVE ROBOT ARCHITECTURE

As the complexity of a task grows, so do the software complexities necessary to process sensory information and to control actions purposefully. Development and maintenance of complex or large-scale software systems can benefit from domain-specific guidelines that promote code reuse and integration through software agents. Information processing in our humanoid robot ISAC (Intelligent SoftArm Control) is integrated into a multiagent-based software architecture based on the Intelligent Machine Architecture (IMA) [9]. IMA is designed to provide guidelines for modular design and allows for the development of subsystems from perception modeling to behavior control through the collections of IMA agents and associated memories, as shown in Figure 3 [10].

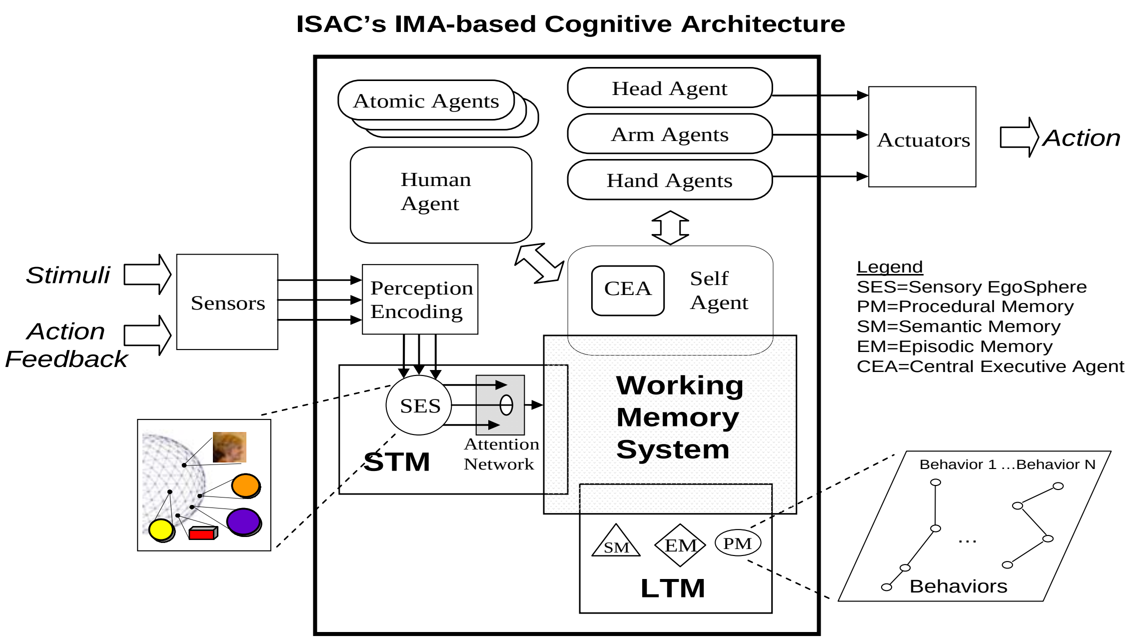


Figure 3. Multiagent-based cognitive robot architecture

ISAC's memory structure is divided into three classes: Short-term memory (STM), long-term memory (LTM), and the working memory system (WMS). STM holds sensory information of the current environment. LTM holds learned behaviors, semantic knowledge, and past experiences. WMS holds task-specific information called “chunks” and streamlines the information flow to the cognitive processes during the task execution. STM is implemented using a sparse sensory data structure called the Sensory EgoSphere (SES). SES, inspired by the egosphere concept defined by Albus [11], serves as a spatio-temporal STM for a robot [12]. LTM stores information such as *skills learned* and *experiences gained* for future retrieval.

4. COGNITIVE CONTROL AND THE WORKING MEMORY SYSTEM

Intelligent robots should possess the ability to act appropriately in highly dynamic and complex environments involving attention, learning, and memory. Such robots need to exhibit a higher form of cognitive ability called *cognitive control*.

4.1 Cognitive Control

Stated differently, cognitive control in humans is the ability to “consciously manipulate thoughts and behaviors using attention to deal with conflicting goals and demands” [13]. As levels of human behavioral processes range from reactive to full deliberation, cognitive control must also be able to switch between these levels to cope with the demand of task and performance, particularly in novel situations. According to a number of cognitive psychologists, cognitive control in human is performed through the working memory in the pre-frontal cortex (PFC) [14][15]. We will describe how we are realizing cognitive control for ISAC.

4.2 Central Executive and Working Memory System

ISAC’s cognitive control functions are modeled after Baddeley’s psychological human working memory model [16]. In his model, the “central executive” controls two “working memory” systems: the phonological loop and the visuo-spatial sketch pad. Cognitive control in ISAC is currently implemented using STM, LTM, the Attention Network and WMS. STM handles sensor-based percepts that are assigned the *focus of attention* or *gating* by the Attention Network [17]. Perceived sensory inputs that have a high emotional salience, i.e. task-related chunks, will cause the Attention Network to post them to WMS.

Working memory “represents a limited-capacity store for retaining information over the short term and for performing mental operations on the contents of this store. This type of memory system is said to be closely tied to “task learning and execution” [18].

Inspired by this, we have implemented the working memory structure into ISAC to provide the embodiment necessary for exploring the critical issues of task execution and learning [18]. Our hypothesis is that this integration will lead to a more complex, but realistic robotic learning system involving perceptual systems, actuators, reasoning, attention, emotion, and short- and long-term memory structures [19].

4.3 Control Architecture

The architecture shown in Figure 4 was implemented for the experiment described in Sect. 5. The architecture centers around WMS, which enables ISAC to make appropriate connections between task commands and focus of attention.

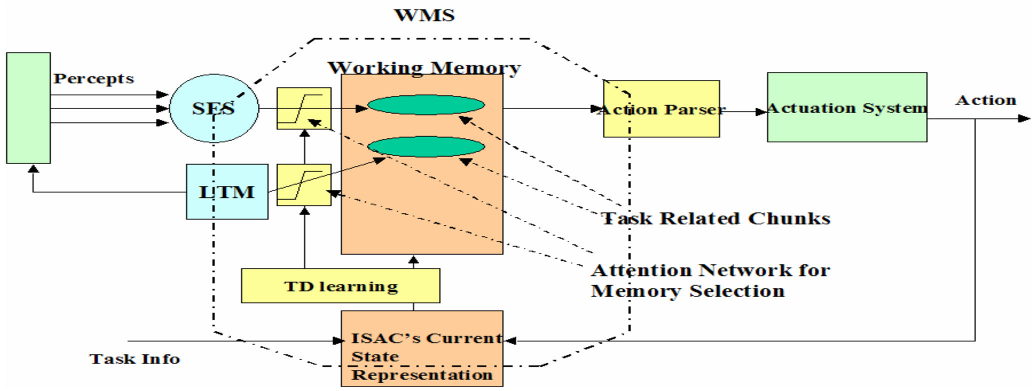


Figure 4. Control architecture used for behavior learning and generation experiment

This architecture integrates the perceptual system, motion generation (actuation) system, SES, and LTM. As shown in Figure 4, SES and LTM act as virtual components of WMS due to the fact that WMS can be comprised of information found in either or both. Part of WMS is the system's state representation used to monitor the task-related chunks kept in the working memory. TD learning [20] allows the system, over time, to load chunks that it expects the most reward for, given the current state. There is also an action parser that identifies whether a chunk is a behavioral chunk or a perceptual chunk and then sends the appropriate command (such as "perform action X on percept Y") to the actuation system.

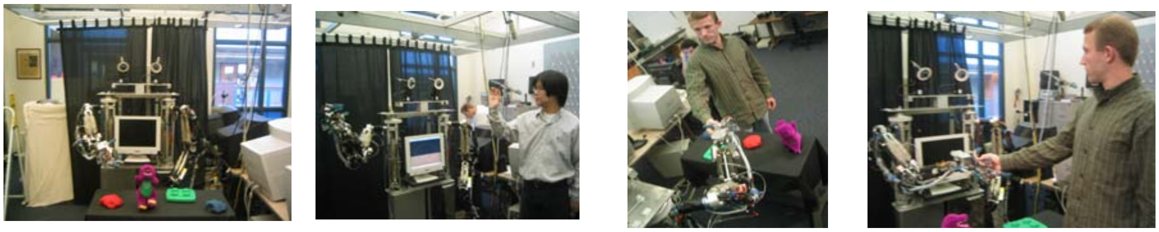
5. BEHAVIOR LEARNING AND GENERATION EXPERIMENT

An experiment was designed to demonstrate the effectiveness of system performance using portions of the cognitive robot architecture.

5.1 Experiment

Steps for this experiment are as follows:

1. ISAC is given certain initial knowledge [Appendix 9.3]:
   a. ISAC's perceptual system is trained to recognize specific objects. The information is stored in the semantic memory section of LTM.
   b. Using the *verbs and adverbs* algorithm [21, Appendix 9.2], ISAC is taught a small set of motion behaviors including how to *reach*, *wave*, and *handshake* (i.e., *reach* and *grab*) (Figure 5). Information is stored in the procedural memory section of LTM.

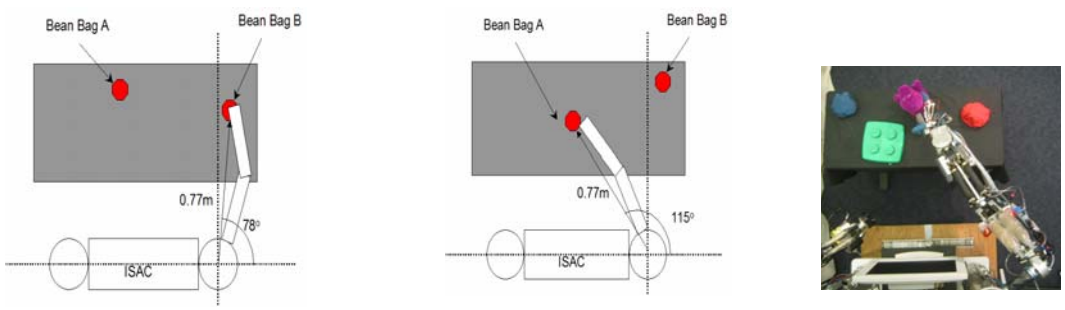


Reach *Wave* *Handshake (over and front views)*
Figure 5. Sample configurations for reaching and actual experiment lab view

2. Two bean bags are placed on a table as shown in Figure 6.a.
3. ISAC is asked to "reach to the bean bag", though the bean bag is not specified.
4. ISAC's perceptual system will recognize the bean bags and post the information to SES.
5. WMS will focus on "chunks" of information necessary for accomplishing the task.

6. A reward is given based upon how well the action is completed.
7. Over time, ISAC learns the appropriate chunks from the SES and LTM to focus on.
8. Once ISAC has demonstrated that it has learned the most appropriate chunks to load into WMS, bean bags are rearranged (Figure 6.b) and ISAC will again be given the command “reach to the bean bag”.

When bean bags are rearranged, ISAC should not necessarily reach to the same bean bag as before but should choose the *bean bag percept* from SES that is the most appropriate. For this task the most appropriate bean bag is the nearest one. The combination of percept and behavior, or “chunks”, will be loaded into the working memory and used to execute the action.



(a) *(b)* *(c)*

Figure 6. Sample configurations for reaching and actual experiment view

5.2 System Performance

Reward was given based on the following three criteria:

1. What was the degree of success for the behavior WMS chose to load?
2. How well did the object chosen by WMS meet the task criteria? e.g., as long as ISAC focused on a bean bag for the “reach to bean bag” task, a reward was given.
3. How well was ISAC able to act upon the object? e.g., in this experiment, could ISAC reach the bean bag?

In order to measure Reward Criterion #3, the reward was given based on the inverse proportion of the distance from ISAC’s hand to the object. Reward Criteria #1 and #2 gave a discrete positive valued reward if the system chose appropriately. No preference (i.e., reward of 0) was the result if the system did not choose correctly. The values for the overall reward typically fell in the range of 0 – 400. Since it was desired to give negative reward to the system when it did not act appropriately, a negative weighting factor of –200 was added to the final reward to “tilt” the low values into the negative range.

Note that using these reward criteria, it is possible to incorrectly reward the system for performing the task in less than an optimal manner. For example, if the system were to *handshake* or *wave* the appropriate bean bag and if this action happened to bring the hand very close to the bean bag, then the system would receive a positive reward. In order to avoid this undesirable situation, more rules or knowledge are needed.

5.2.1 Initial Trials

Initial trials were performed in simulation to speed-up initial testing phase of the system. The simulation was set-up to remove the time-bottleneck of generating and performing motions. If the simulation attempted to act on an object within ISAC’s workspace, it was assumed that ISAC was able to reach to the object (Reward Criterion 3).

The action taken was determined by what WMS currently believed was the best choice. In other words the action that WMS believed would yield the greatest reward. This system also contained an exploration percentage, specified as a part of initial knowledge that determined the percentage of trials that WMS chose a new or different action. This enabled WMS to always continue learning and exploring.

During the initial trials, the simulation was not allowed to choose the same action more than twice. This constraint enabled a much quicker simulation time. Once the system finished exploration, the system was restarted with the learned information and given the task to "*reach* to the bean bag". For each arrangement (Figure 6.a and 6.b) the system chose appropriately to reach towards the correct bean bag, i.e. the nearest one. Table 1 shows the contents of ISAC's short-term and long-term memory systems during the training portion of the simulation.

SES	LTM
1. bean bag: location = (Figure 6.b), type = A	1. *reach*
2. bean bag: location = (Figure 6.a), type = B	2. *handshake*
	3. *wave*

Table 1. Memory contents during simulation training

	Working Memory Contents			
Trial #:	1	2	3	4
Chunk 1	bean bag A	bean bag B	*wave*	*handshake*
Chunk 2	*reach*	bean bag A	bean bag B	bean bag A
Random:	NA	*handshake*	NA	NA
Reward:	203.4	-20.5	-197.7	2.3

Table 2. Working memory contents during simulation training

In these trials, WMS was allowed to choose two "chunks" from the short- and long-term memory systems to accomplish the task. However, the working memory was not restricted to choosing exactly one object and one behavior. If the working memory chose to focus on two objects or two behaviors, then respectively a behavior or object was chosen at random. This ensured that an action was still performed. The reasoning behind this was so that the system did not learn to simply choose combinations that lead to no reward, a situation that could be preferred if WMS was consistently getting negative reward for its choices. Table 2 shows samples of the contents in the working memory in these trials.

To evaluate system performance further, a third task was developed. For this task ISAC was again given the command to "*reach* to the red bag", however this time the *reach* behavior was deleted from the initial knowledge limiting the behavior choices to *handshake* and *wave*. The working memory had to choose the *next best* behavior. For each of the arrangements shown previously (Figures 6.a and 6.b), WMS chose to perform the *handshake* behavior. This behavior was chosen because it allowed the arm to get closest (Reward Criterion 3) to the bean bag (Reward Criterion 2) and thus best accomplished the goal task.

5.2.2 Trials on ISAC

After the initial training, ISAC was allowed to perform the generated motions (Figure 6.c). Two new objects (a green Lego toy, and a purple Barney doll) were added to the table, at random positions. ISAC's vision system was trained (Step 1 in Sect. 5.1) to recognize each new object and recorded the type of object as well as some simple descriptive information (color=green, purple; toy type=Barney doll, Lego). ISAC was given tasks (Step 3) such as "*reach* to the bean bag" or "*reach* to the toy". Each of these tasks did not specify to which bean bag or toy ISAC was to *reach*. ISAC recognized the objects (Step 4). WMS focused on "chunks" of information from the SES and LTM in order to accomplish the task (Step 5). ISAC was allowed to explore the space of possible actions receiving reward each time (Steps 6 and 7). After this was accomplished, the objects were rearranged in a variety of different positions (Step 8) and ISAC was given a command. The results (set of 20 commands) were that ISAC successfully performed the correct action on the nearest (easiest to reach) requested object.

5.2.3 Performance Evaluation

For this system to properly choose the correct set of chunks to focus on, the system currently has to explore all the possibilities during training. Attempting to explore all possibilities will obviously lead to a combinatorial explosion if a large number of behaviors or percepts are added to the system. In order for this system to operate properly in complex tasks there needs to be some mechanism in place to aid the working memory in limiting its search to only the most relevant chunks. To achieve this goal, we are now implementing a cognitive agent called the Central Executive Agent (CEA) (see Figure 3). CEA will work closely with the working memory system to provide such an intelligent selection mechanism. An initial version of CEA has been tested in simulation to evaluate ISAC's decision-making process for *simulation-based response* using a set of initial knowledge about acoustic stimuli (i.e., alarm sound and a musical piece) and learning to choose a proper action among conflicting goals (i.e., to continue the task at hand or to switch to another [22].

6. CONCLUSIONS

During the past decade, we have seen major advances in the integration of sensor technologies, artificial intelligence, and machine learning into a variety of system design and operation. A next challenge for control engineers will be the integration of human-like cognitive control into system design and operation.

This paper described our efforts to develop the next generation of robots with robust sensorimotor intelligence using a multiagent-based robot control architecture and a biologically inspired intelligent control. Simulation and real-time experiments were conducted to demonstrate the effectiveness of our design. The next step is to integrate a cognitive control experiment involving the Central Executive Agent.

ACKNOWLEDGEMENTS

Special thanks to Palis Ratanaswasd, Flo Fottrell, and the graduate students at the Center for Intelligent Systems. 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] Technical Committee on Intelligent Control, IEEE Control Systems Society, http://robotics.ee.nus.edu.sg/tcic/

[2] M.S. Gazzaniga, Cognitive Neuroscience. *The Biology of the Mind*, 2nd ed., Norton & Co., New York, 2002, p. 615.

[3] J.R. Stroop, "Studies of interference in serial verbal reactions," Jo. of Experimental Psychology, 18, 1935, pp. 643-661.

[4] M.M. Botvinick, et. al, "Conflict monitoring and cognitive control," Psychological Review, 108(3), 2001, pp. 624-652.

[5] J.R. Anderson, et al., "An integrated theory of the mind", (ACT-R), Psychological Review, 111(4), 2004, pp. 1036-1060.

[6] J.E. Laird, A. Newell, and P.S. Rosenbloom, "SOAR: An architecture for general intelligence", Artificial Intelligence, 33(1), 1987, p. 1-64.

[7] J.M. Mendel and K.S. Fu, eds., *Adaptive, Learning and Pattern Recognition Systems: Theory and applications*, Academic Press, New York, 1970.

[8] E.K. Miller, C.A. Erickson, and R. Desimone, "Neural mechanisms of visual working memory in prefrontal cortex of the macaque", Jo. of Neuroscience,16, 1996, pp. 5154-56.
[9] R.T. Pack, D.M. Wilkes, and K. Kawamura, "A software architecture for integrated service robot development," Proc. of IEEE Systems, Man and Cybernetics, Oct. 1997, pp. 3774-3779.
[10] K. Kawamura, R.A. Peters II, R. Bodenheimer, N. Sarkar, J. Park, A. Spratley, and K. A. Hambuchen, "Multiagent-based cognitive robot architecture and its realization," Int'l Jo. of Humanoid Robotics, 1(1), 2004, pp. 65-93.
[11] J. S. Albus, "Outline for a theory of intelligence," IEEE Trans. on Systems, Man, and Cybernetics, 21(3), 1991, pp.473-509.
[12] R.A. Peters II, K.A. Hambuchen, K. Kawamura, and D.M. Wilkes, "The sensory egosphere as a short-term memory for humanoids," Proc. of the IEEE-RAS Int'l Conf. on Humanoid Robots, Waseda University, Tokyo, Nov.22-24, 2001, pp.451-459.
[13] M. Minsky, *The Society of Mind*, NY: Simon & Schuster, 1985.
[14] T. S. Braver and J. D. Cohen, "On the control of control: The role of dopamine in regulating prefrontal function and working memory," in S. Monsell & J. Driver, eds., *Control of Cognitive Processes: Attention and Performance XVIII*, Cambridge, MA: MIT Press, 2000, pp. 713-738.
[15] S. Funahashi and K. Kubota, "Working memory and prefrontal cortex", Neuroscience Research, 21, 1994, pp. 1-11.
[16] A.D. Baddeley, *Working Memory*, Oxford: Clarendon Press, 1986.
[17] K.A. Hambuchen, Multi-Modal Attention and Binding using a Sensory EgoSphere, Ph.D. Dissertation, Vanderbilt, Nashville, TN, May 2004.
[18] Phillips, J. L. and D. Noelle, "A Biologically Inspired Working Memory Framework for Robots", Proc. of the 27th Annual Conf. of the Cognitive Science Society, 2005, pp. 1750-1755.
[19] M. Skubic, D. Noelle, M. Wilkes, K. Kawamura, and J.M. Keller, "A biologically inspired adaptive working memory for robots," AAAI Fall Workshop on the Intersection of Cognitive Science and Robotics: From interfaces to intelligence, Washington DC, October 2004.
[20] S. Russel and P. Norvig, *Artificial Intelligence: A modern approach*, 2nd ed., Pearson Education Inc. 2003.
[21] C. Rose, M.F. Cohen, and B. Bodenheimer, "Verbs and adverbs: Multidimensional motion interpolation," IEEE Computer Graphics & Applications, 18(5), 1998, pp. 32-40.
[22] P. Ratanaswasd, C. Garber, and A. Lauf, "Situation-Based Stimuli Response in a Humanoid Robot", accepted, the Fifth International Conference on Development and Learning, Bloomington, IN., May 31-June 3, 2006.
[23] T.S. Braver and D.M. Barch, "A theory of cognitive control, aging cognition, and neuromodulation," Neuroscience and Behavioral Reviews 26, St. Louis: Elsevier Science Ltd., 202, pp. 809-817.

APPENDICES

1. *Stroop Test*

In the traditional Stroop Test, the subject must name the color in which a word is written [3]. In this task, the context must be actively maintained to direct attentional resources and response selection. Competition among responses arises "when the appropriate response is one that is relatively infrequent or when the inappropriate response is dominant." [23] For instance, when the word *green* is presented in the *color red*, the automatic response would be to say "red", but the appropriate response is "green". Cognitive control plays a major role in the test.

2. Verbs and Adverbs

Verbs and adverbs is a motion interpolation algorithm originally developed for computer graphics [22]. This technique has been modified for applications on ISAC [11]. In short, the algorithm interpolates motions across different motions spaces. For example, a motion *handshake* is termed a *verb* with parameters of the handshaking such as the direction of the motion serving as the *adverbs*. These *adverbs* are represented by a range of values from (min, max). A *verb* can have any number of *adverbs*. *Handshaking* has one *adverb*: the direction of the motion. *Reaching* has three *adverbs*: the polar coordinates of the object to be reached towards. *Waving* has two *adverbs*: the height and direction of the motion. The range of values for each *adverb* represents the boundaries for that aspect of the motion. If ISAC wanted to shake the hand of someone standing in the middle of ISAC's workspace, ISAC would interpolate a new *handshake* motion with the *adverb* value (min + max)/2. When given a command, ISAC determines the appropriate *adverb* values based on the current perceptual information, typically the location of the object. For safety, ISAC does not attempt to extrapolate motions outside of the *adverb* space.

3. Initial Knowledge used for the Experiment

Initial knowledge for this experiment is the knowledge that ISAC needs for the experiment. This knowledge could have been programmed into ISAC or ISAC could have been shown/told this knowledge at an earlier time. Prior to this experiment a handful of distinctly colored objects were shown to ISAC. These objects were then labeled with semantic descriptors denoting the type, color, and texture of each object, the algorithm to identify the object, and any necessary parameters that algorithm requires to identify the object. The objects shown to ISAC were a variety of colored bean bags, a green lego toy, and a purple doll. All of this information was stored in ISAC's semantic LTM.

ISAC was also taught how to perform simple behaviors (*reach*, *handshake*, and *wave)* through teleoperation. Each behavior was given a name and the appropriate *adverb* ranges [Appendix 9.2] and stored in ISAC's procedural LTM.