Development of a HAM-style Cognitive Robot

Kazuhiko Kawamura
Professor and Director
Center for Intelligent Systems / Vanderbilt University
Nashville, Tennessee  37235-0131 USA
http://eecs.vanderbilt.edu/CIS

Abstract. A robotic-aid system could be more effective if the system were intelligent enough to understand human needs and adapt its behaviors accordingly. This paper presents our efforts to realize such a robot that we call a cognitive robot through a multi-agent-based cognitive robot architecture with three distinctive memory systems. Several applications illustrate how the system interacts with humans.

Keywords: Cognitive robot, human-robot interaction, robotic-aid system, multi-agent-based architecture

I. INTRODUCTION

Robotics has evolved from the industrial robots in the 1960s to nontraditional robots for surgery and rehabilitation in the 2000s. One class of robot that is gaining popularity is the anthropomorphic or humanoid robot. Starting in 1990s, the Cognitive Robotics Laboratory of Vanderbilt University has been developing a humanoid robot called the Intelligent Soft-Arm Control (ISAC) [1]. Originally designed to assist the physically disabled [2] (Figure 1), ISAC gradually became a general-purpose humanoid robot to work with a human as a partner or an assistant [3] (Figure 2).

In designing a robotic-aid system, the need is to maximize the adaptability of the human-robot interaction and to produce a human-friendly interface with the robot. Our goal is to develop a personalized robotic system to achieve just such a human-friendly symbiotic system. Personalized means that the robot behaves differently depending on the state of the human and the robot’s emotions. For example, if the user is a small child, the robot may decide to move more slowly than normal in order not to frighten the child. Or, you expect high quality service from a waiter who knows you well, so the robotic waiter will serve you to satisfy your needs. Hence, we believe that the development of a personalized robotic-aid system is a valuable step toward a human-friendly mechatronic (HAM) system.

In order to realize such a robot, we developed a multi-agent-based cognitive architecture incorporating several types of memory structures for parallel, distributed robot control [4].

In this paper we report recent progress in developing two high-level agents, the Human Agent and the Self Agent, plus memory structures. The Human Agent (HA) is the humanoid’s internal representation of the human. The HA includes information concerning the location, activity, and state of the human, as determined through observations and conversations. The Self Agent (SA) is the humanoid’s internal model of itself. The SA provides the system with a sense of self-awareness concerning the performance of the hardware, as well as the progress and effectiveness of tasks and behaviors.

Our approach to robot memory structure is through short- and long-term memories called the Sensory EgoSphere (SES) and the Procedural/Declarative/Episodic Memory, respectively and an adaptive working memory system (WMS).
II. DESIGN OF PERSONALIZED HUMAN-ROBOT INTERACTION

2.1 Categorizing Users

The first step of developing personalized robotic-aid systems is to categorize users based on their physical and psychological conditions. Physical conditions include a user’s height, weight and age. For disabled people, various conditions will exist depending on their body conditions. Psychological conditions includes preferred working speed and the willingness to work with the robot.

2.2 Framework for Human-Robot Interaction

In the CIS, our philosophy for software design for a humanoid is to integrate both the human and the robot in a unified multi-agent based framework [4]. Thus, we group aspects of Human-Robot Interaction (HRI) into the three categories:

- **Physical**: structure or body, e.g., physical features, manipulation capabilities
- **Sensor**: channels used to gain information e.g., voice, vision
- **Cognitive**: internal working of the system, e.g., human (mind and affective state) and humanoid (reasoning, communication of its intention).

Because it can be difficult to determine cognitive aspects of humans consistently, we limit our cases where both humans and humanoid intend to achieve a common goal.

We are also interested in giving the humanoid its own emotional or affective module to make HRI more socially pleasant [5]. ISAC is equipped with sensors (cameras, microphones, infrared sensors) for capturing communication modes including face detection, finger pointing, etc. An infrared motion detector provides ISAC with a means of sensing human presence. We use MS Speech engines for detecting human speech and use a sound-localization system [6].

![Figure 3. IMA architecture and Memory Structure](image)

Likewise, we are developing techniques for ISAC to give feedback to people through speech and gestures. We are also developing the use of a visual SES display projected on ISAC’s monitor. The interface is based on multi-agent-based architecture called *IMA*, the Intelligent Machine Architecture, developed at Vanderbilt [4]. Figure 3 illustrates the overall IMA agent structure with the memory structure.

III. HUMAN AGENT FOR PERSONALIZED HUMAN-ROBOT INTERACTION

3.1 The Human Agent

The Human Agent comprises a set of agents that detect and keep track of human features and estimate the intentions of a person within the current task context. It estimates the current state of people interacting with the robot based on observations and from explicit interactions (Figure 4) [7]. The HA processes two types of human intentions, i.e., expressed and inferred.

The Human Agent’s assessment of how to interact is passed on to the SA. The SA interprets the context of its own current state, e.g., current intention, status, tasks, etc. This processing guides ISAC in the selection of socially appropriate behaviors that lead towards the ultimate goal of completing tasks with (or for) humans.

![Figure 4. Human agent and associated atomic agents](image)

Figure 5 shows the model of the levels of interaction engagement, represented by a numerical value, that we have developed as the basis for modeling social interaction. These levels progress the robot from a state of no interaction to an ultimate goal of completing a task with (or for) a person. Level 1, Solitude, corresponds to when ISAC does not detect anyone in the environment. In this situation, ISAC may choose actions to actively attract people with whom to interact. Level 2, Awareness of People, corresponds to a stage, often short time, when ISAC is aware of people around it, and has not interacted with them. Level
3. Acknowledgement, is the phase when the robot actively acknowledges the presence of a person. This is performed if the person is approaching ISAC for the first time or if the person is interrupting an ongoing interaction. Level 4, Active Engagement, represents that stage of active interaction.

![Figure 5. Interaction levels within Human Agent](image)

3.2 Demonstration in situation-based acknowledgment

In this demo, ISAC processes the intentions of the human, resolves them with its own intentions and abilities, and communicates to the person if there is a problem with the request. The scenario begins as a person approaches ISAC and gains its attention. The Human Agent determines that the person has an intention to interact with ISAC. If ISAC is unoccupied at the time, ISAC begins its interaction behaviors by turning toward the person and initiating a greeting and identification sequence. Once interaction is established, ISAC begins a social dialog. After greeting, ISAC may respond to a person's task request (an intention for ISAC to do something). If a second person approaches ISAC and attempts to gain its attention, the Human Agent will notify the Self Agent that there is a new person with a pending intention (Figure 6a). The Self Agent must then resolve the current human intention with its own current intention. If the second human intention is not of sufficient priority to override ISAC's current task, then ISAC will pause its current interaction, turn to the interrupter, and apologize for being busy (Figure 6b).

![Figure 6. (a) ISAC responding to an interruption; (b) Role of intention processing during interruption](image)

IV. SELF AGENT FOR COGNITIVE CONTROL

4.1 HAM-style cognitive robot

HAM-style robots could be categorized as (1) prosthetic hands/arms/legs, (2) haptic-based robotic aid systems (see Figure 1), and (3) robotic assistants/partners such as robot helpers or nurses. HAM-style cognitive robots described in this paper fall into the last category. Our efforts, thus, may be called the cognitive approach to HAM-style robot development.

4.2 Cognitive and Conscious machines

As we enter the new century, a considerable amount of research is being conducted in the area of cognition and consciousness in humans and machines. An increasing number of scientists today agree that cognitive or conscious machines is not only possible but also inevitable for that to happen [8][9][10][11][12].

Before fully explaining what machine cognition or consciousness is, it may be advantageous to build a machine which can demonstrate key functions of a cognitive agent so that many questions will be clarified. For example, we are working on the Self Agent which uses emotions, attentions and cognitive control to deal with new situations.

4.3 Self Agent

The Self Agent (SA) is responsible for ISAC’s cognitive activities ranging from sensor signal monitoring and task monitoring to decision-making. The SA integrates failure information from sensors and maintains information about the task-level status of the humanoid. Its cognitive aspects include recognition of human intention and selection of appropriate actions by forming plans and activating them. The current structure and the functions of the Self Agent are provided in Figure 7. Conflict resolution is currently handled based on the relative priority of the conflicting actions.
ISAC's memory structure is divided into three classes: Short-Term Memory (STM), Long-Term Memory (LTM), and the Working Memory System (WMS). The STM holds information about the current environment while the LTM holds learned behaviors, semantic knowledge, and past experience, i.e., episodes. The WMS holds task-specific STM and LTM information and streamlines the information flow to the cognitive processes during the task.

5.1 Short-Term Memory: Sensory EgoSphere

Currently, we are using a structure called the Sensory EgoSphere (SES) to hold STM data. The SES is a data structure inspired by the egosphere concept as defined by Albus [16] and serves as a spatio-temporal short-term memory for a robot. The SES is structured as a geodesic sphere that is centered at a robot's origin and is indexed by azimuth and elevation [17].

The objective of the SES is to temporarily store exteroceptive sensory information produced by the sensory processing modules operating on the robot. Each vertex of the geodesic sphere can contain a database node detailing a detected stimulus at the corresponding angle (Figure 9).

5.2 Long-Term Memory: Procedural, Episodic, and Declarative Memories

LTM is divided into three types: Procedural Memory, Episodic Memory, and Declarative Memory. Like that in a human brain, the LTM stores information such as skills learned and experiences gained in the long term for future retrieval.

The part of the LTM called the Procedural Memory (PM) holds motion primitives and behaviors needed for movement, such as how to reach to a point. Behaviors are derived using the spatio-temporal Isomap method proposed by Jenkins and Mataric [18]. A short description of how it operates is shown in Figure 10.
Motion data are collected from the teleoperation of ISAC. The motion streams collected are then segmented into a set of motion primitives. The central idea in the derivation of behaviors from motion segments is to discover the spatio-temporal structure of a motion stream. This structure can be estimated by extending a nonlinear dimension reduction method called Isomap [19] to handle motion data. Spatio-temporal Isomap dimension reduction, clustering and interpolation methods are applied to the motion segments to produce Motion Primitives (Figure 10). Behaviors are formed by further application of the spatio-temporal Isomap method and linking Motion Primitives with transition probabilities [20].

![Figure 10. Derivation of Procedural Memory through human-guided motion stream](image)

Motion skills for each behavior must be interpolated in order to be used in specific situations. The interpolation method we are using is the Verbs and Adverbs method developed in [21]. This technique describes a motion (verb) in terms of its parameters (adverbs) that allows ISAC to generate a new movement based on the similarity of stored motions.

Currently under development, the Episodic Memory (EM) system will hold records of past experiences in a time-indexed format. The purpose of the EM is to allow the cognitive process to trigger new action, selection, and execution. The Working Memory System records its own contents with reward information for the duration of a single task and posts this information to a LTM data unit.

The Declarative Memory (DM) currently is a data structure about objects in the environment. In the future, we plan to expand to include semantic knowledge.

VI. INTEGRATED COGNITIVE CONTROL EXPERIMENT

We have designed an integrated cognitive system experiment using the CEA, attention, emotion and the adaptive working memory system as follows:
1. ISAC is trained to learn specific object using voice, vision and attention (Learning by association)
2. ISAC is asked to point to one of the learned objects (Use of short-term memory of the object and long-term procedural memory)
3. ISAC is asked to visually track the object held by a human (Color tracking)
4. A person enters the room and yells "Fire!" ISAC using attention, emotion and cognitive control, suspend the current tracking task and warn everyone to exit the room (Cognitive control)

Steps 1-3 have been implemented and presented elsewhere [22][23]. In Step 4, ISAC's cognitive control must
- Pay attention to new stimulus
- Use emotion to activate the episodic memory
- Use EM to activate cognitive control.

This cognitive control experiment is being done through integrating the WMS and cognitive control with the various IMA agents (Figure 11).

![Figure 11. Cognitive control experiment](image)

This experiment tries to demonstrate that, "The artificial cognitive machine is not governed by any programs and therefore will not execute any preprogrammed decision commands like the IF-THEN ones" [p.216, 8].

VII. CONCLUSION

In order to realize a personalized robotic-aid system with our humanoid robot, ISAC, we developed a multi-agent-based architecture. In particular, we developed two high level agents, the Human Agent and the Self Agent. The Human database within the Human Agent stores personalized data, and other agents within the Human Agent are used to recognize each person’s intention and state. The Self Agent is responsible for performing the behaviors requested by the Human Agent according to the user’s state. We believe that this approach is effective for achieving personalized human-robot interaction, resulting in a more HAM-like robot.
ACKNOWLEDGEMENTS

The author would like to thank members of the Cognitive Robotics Lab, past and present, and Professor David Noelle for their contributions.

REFERENCES


