# Task-Learning Policies for Collaborative Task Solving in Human-Robot Interaction

Hae Won Park
Human-Automation Systems Lab
Georgia Institute of Technology
85 5th st NW, Atlanta GA, 30308
haewon.park@gatech.edu

## ABSTRACT

The objective of this doctoral research is to design multi-modal task-learning policies for a robotic system that targets the exchange of task rules between humans and robots. This objective is achieved through a collaborative task application during human-robot interaction where the two partners learn a task from each other and accomplish a shared goal. As a first step, a method to model human-action primitives using a pattern-recognition technique is presented. Next, algorithms are developed to generate turn-taking strategies in response to human task behaviors. The contribution of this work is in engaging robots with humans in collaborative play task by modeling statistical patterns of play behaviors and reusing previously learned knowledge to reduce the decision process. Here, results of previous work are presented, and remaining works including deploying a physically embodied agent and developing an evaluation platform are outlined.

## Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics; I.2.6 [**Artificial Intelligence**]: Learning; I.2.10 [**Artificial Intelligence**]: Vision and Scene Understanding

## Keywords

Human-robot interaction, Collaborative task learning and solving

## 1. INTRODUCTION

The need for robots being able to socially interact with humans is becoming greater as more applications emerge that require interactions between human and robot. Robots that interact with humans or other physical agents by following social behaviors and rules attached to their roles are defined as social robots [2]. In 2005, Feil-Seifer and Matarić [3] introduced the term *socially assistive robots* (SAR) that extends on this definition. Traditionally, assistive robots referred to mechanical actuator devices built to assist people with disabilities or help with physical rehabilitations. Compared to these devices, examples of SAR include robots that create cognitive bonding with the user by responding to the user's touch and sound [9]. The focus of this doctoral research is on SAR systems that possess the ability to engage in collaborative activity, which in turn promote a subject's social skills and mediate the process of transferring the subject's learned knowledge for interaction with other people. As a social mediator, the child-like humanoid Kaspar has shown potential in encouraging autistic children to participate in an imitation play [6]. In this work, the child and the therapist take turns imitating Kaspar's expressions. Some subjects, after observing the robot play with a tambourine, start to mimic the action. The authors also mention how robots generate a high degree of motivation and engagement in subjects, including those who were unwilling to interact with human therapists. Their subjects showed positive social behaviors, such as touching, vocalizing, and smiling at the robot. The result of the study promises that repeated exposure to interactive robots can effectively enhance the social ability of the subjects. Compared to this work where the robot was remotely controlled by the therapist, the goal of this research is to develop a robotic system that could observe, understand the subject's play, and autonomously generate its own turn-taking behavior.

In the following sections, the key research questions are posed, followed by discussions on current accomplishments and approaches for the remaining work.

## 2. RESEARCH QUESTIONS

Endowing robots with learning capability is one of the active research areas in robotics. Learning from demonstration (LfD) can greatly reduce time for computing task constraints or planning manipulation trajectories. In the studies where LfD was coupled with social machines, it was shown that social guidance promotes the robot's learning process [8]. By building on these previous efforts, this doctoral research focuses on implementing a task-learning behavior on a robotic platform where the robot (1) observes and interprets human motions, (2) deduces the underlying objective, (3) generates an appropriate response, and (4) takes turns with the human partner to accomplish a task. Play tasks have been chosen as the application domain in order to evaluate the proposed task-learning system. Specifically, the following questions are posed and answered through this research.

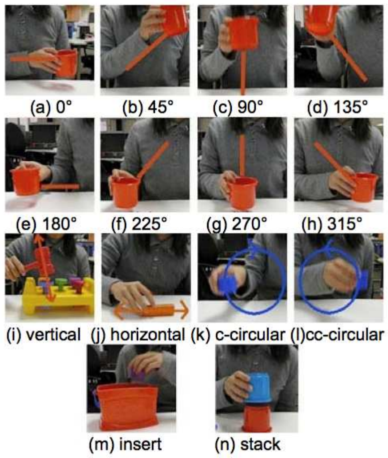


**Figure 1: The most frequently observed low-level motions are defined as play-task primitives.**

1. Given a task objective, how should a task behavior be effectively modeled and recognized?
2. What are the methods to autonomously select and perceive task-learning cues (TLCs) in a scene?
3. In what way should the TLCs be combined to generate a transferable task-learning policy (TLP)?
4. What metrics should be used to measure learning performance, i.e., task knowledge transfer, of a robotic agent?

The following two sections report the approaches and findings of the research questions 1 and 3. The future work focuses on improving the previous results and providing answers to the remaining questions.

## 3. TASK-BEHAVIOR MODELING

Baranek *et al.* list in their work a subset of toy manipulations that could be used towards screening a child's developmental stage [1]. Based on the list, a study has been conducted with public sources from the web to identify the kinds of basic motions that form these manipulations when people interact with various types of toys. With regards to a robotic partner, these basic motions are further defined as *play-task primitives*. The fourteen most distinct primitives found from the study are the renditions of the behaviors in the list of Baranek *et al.* (Figure 1).

Much of the prior literature refers to hidden Markov model (HMM) as a very effective gesture pattern-recognition algorithm [7]. The HMM provides a probabilistic framework for modeling a time series of multivariate observations. The power of the algorithm comes from the characteristic that defines a Markov process. In a Markov process, the conditional probability distribution of the future event only depends on the current event and not on the sequence of predecessor events. In modeling play actions, it might not be valid to assume that the entire play sequence has a Markov property. However, the assumption is still useful when considering the constant changes in motion gradients and object appearances that form the play motions. The contribution of this approach is in decomposing a large action to temporally-sequenced primitives with a first order Markov process, thereby providing recognition of unpredicted behaviors.

The task-primitive recognition algorithm is depicted in Figure 2. The features used to model the primitives are motion gradients and object spreads. What defines the movement of an object is the directional difference in the object's center of mass in adjacent frames. The term motion gradient is used to describe the normalized direction of the task motion. Besides the $x$ and $y$ directional changes of the object movement, the variation in the object size is used to distinguish whether the object has been inserted, stacked, or dropped. Detailed algorithm and results are discussed in [5].

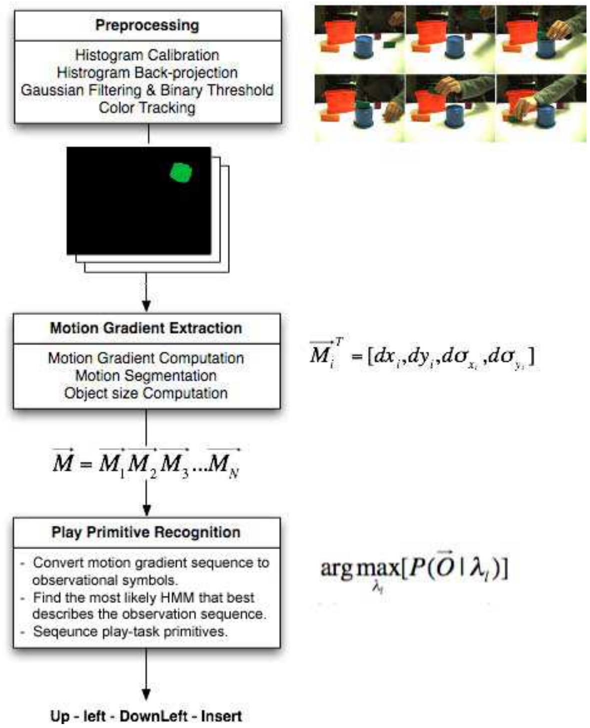


**Figure 2: Steps of training the hidden Markov models and recognizing the sequence of play-task primitives.**

## 4. GENERATING A TASK TURN

Case-based reasoning (CBR) is a concept that solves new problems based on the solutions of similar past problems. By comparing the current task to some past task cases stored in memory, the best solution is retrieved and adapted to the current task. This process allows the system to bypass a long complicated decision process. One of the issues associated with CBR is that complexity of retrieving a data from a database increases exponentially depending on the size of the database. Therefore, efforts have been made to maintain database size under threshold by re-building the database for each new task and grouping cases by task objectives. Also, low-dimensional feature descriptors have been proposed to reduce data size.

The first phase of CBR is acquiring knowledge, i.e., training the database. During this phase, the system observes a play task performed by the subject and generates a case (problem-solution pair) for each turn which is then saved in the database. In the second phase when a new problem is introduce, the most similar past problem and its solution are retrieved from the database. The distance between the two problems $\mathbf{C_i}$ and $\mathbf{C_j}$ is computed as the sum of weighted distances between each feature descriptor:

$$\begin{aligned}\delta(\mathbf{C}_i, \mathbf{C}_j) = \; & \underbrace{w_1 \cdot \delta(\mathbf{p}_{A_i}, \mathbf{p}_{A_j})}_{\text{primary-object SD distance}} + \underbrace{w_2 \cdot \delta(\mathbf{p}_{B_i}, \mathbf{p}_{B_j})}_{\text{secondary-object SD distance}} \\ & + \underbrace{w_3 \cdot \delta(\rho_i, \rho_j)}_{\text{size ratio distance}} + \underbrace{w_4 \cdot \delta(\gamma_{A_i}, \gamma_{A_j})}_{\text{primary-object color distance}} \\ & + \underbrace{w_5 \cdot \delta(\gamma_{B_i}, \gamma_{B_j})}_{\text{secondary-object color distance}}\end{aligned}$$

The weight coefficients $w_1, \cdots, w_5$ provide factor information of the task objective. For example, if $w_1$ and $w_2$, which are the shape descriptors (SD) of the objects, are greater than other weights, the task objective most likely involves shape matching. Next is the reuse step where the retrieved solution is compared to the current task. Using the distance function $\delta(\mathbf{C}_i, \mathbf{C}_j)$, a new solution is generated that adapts to the current scene. During the last phase, the new problem-solution pair is revised and retained in the database. Using double-thresholding variables, the new case is either discarded, saved in the database, or prompted for human input.

The preliminary result of the proposed system in generating a task turn during collaborative play has been published in [4]. The results were compared between the two distance metrics using equally distributed weights and task-adaptive variable weights. An example in Figure 3 shows that when the weights are equally distributed, the system does not cope well in situations when there is no exact match to a retrieved solution in the current scene. The proposed framework was able to deduce a solution within a second with a successful solution rate of 92% using the distance metric with task-adaptive variable weights.

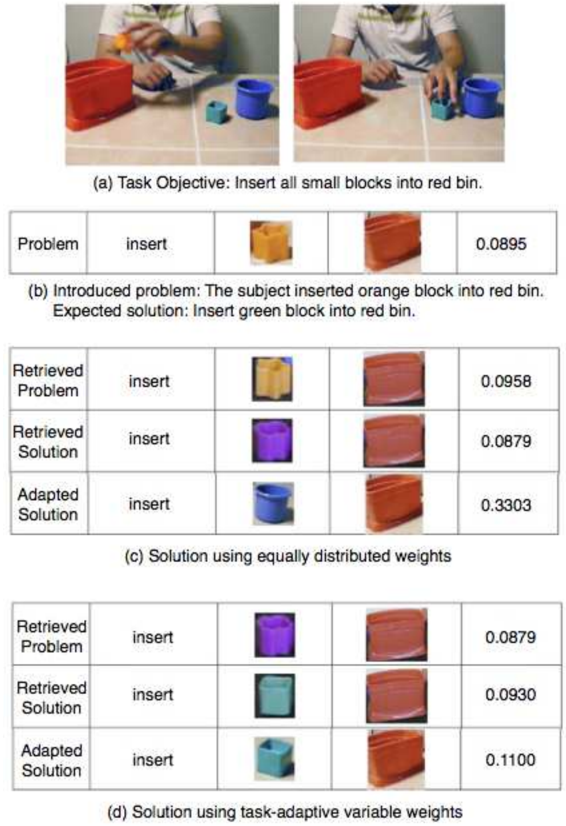


**Figure 3: Example result comparing the solutions deduced by using equally distributed weights and task-adaptive variable weights.**

## 5. REMAINING WORK

The remaining work to achieve the aforementioned goals are detailed in the following sections. In each section, the current status is summarized, and the expected approach to complete each goal is proposed.

### 5.1 Designing a Task-learning Policy

It was discussed in the preliminary research that carefully selected problem descriptors are able to characterize a play task. A new solution was created using the CBR system by retrieving a similar problem and its solution. This approach is capable of generating solutions in a fast pace by reusing the retrieved solutions. However, the current design of the case structure has the following limitations: First, if the problem space in the case-base is sparse, the effectiveness of the retrieval process reduces significantly. Second, the system tends to generate consistent, biased solutions depending on the current case-base. Therefore, the general flow of the task, i.e., the task objective, is difficult to be understood or learned. These limitations will be investigated further and a system that is capable of deducing a task rule will be developed. The previous problem descriptors will serve as task-learning cues (TLC), and the corresponding task-learning policy (TLP) will be deduced in a form of logical equation from observing the patterns of these cues.

### 5.2 Physically Embodied Interactions

The physically embodied interactions, when compared to telepresence robots or an on-screen simulation, produce deeper engagement for the user [10]. Such engagement provides a better understanding of the robot's capabilities and increases the subject's enjoyment in a task-oriented setting. In the previous work, video sequences were collected for training and evaluating the system. In the proposed work, a robotic platform will be used for the evaluation of the presented algorithms.

### 5.3 System-evaluation Application

Whether the robotic agent has achieved a successful task learning should be evaluated through the use of interactive applications. The previous evaluation method using video sequences will be modified to engage both humans and robots. An application will be developed on a tablet computer which will function as a shared workspace. TLP will be demonstrated through a touchscreen interface where the subject and the robot take turns interacting with the objects within the application. The tablet application pro-

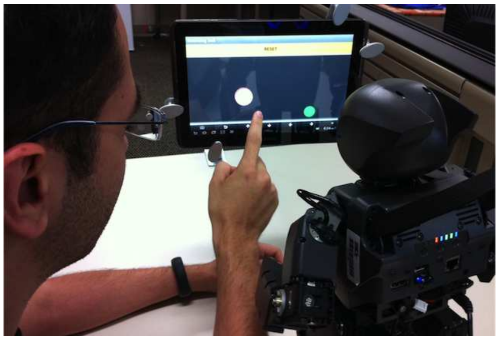

**Figure 4: A touchscreen tablet computer will function as a shared workspace for human-robot collaborative task solving.**

vides versatility in designing different tasks, and using tablet computers reduces expensive kinematic computations of the robot manipulators.

Tablet computers are becoming more available to general populace, and the demand for tablets as assistive-technology platforms is increasing. Therefore, developing a robotic platform that can interact with touchscreen computers provides great potential in deploying robots as companions, therapeutic devices, and playmates.

## 5.4 Subject Testing and Evaluation

In this section, methods for measuring the effectiveness of a task-policy transfer from human to robot are proposed. First, a study will be conducted with multiple video sequences that demonstrate different tasks. Subjects will be invited to predict what the underlying task rules are for each video sequence. The objective of this study is to identify how humans perceive learning cues in a scene, what kinds of visual cues the subjects focus on, and how the subjects link the cues to understand the task rules. The visual cues used for learning the tasks will be added to the previously defined features. The second evaluation will measure how effective the demonstrated task cases are in transferring task policies. The researcher will provide the same set of demonstrations using the tablet application to the subject and the robot. The task rules predicted by the subject will be compared to the task policy generated by the robot. The final experiment will also be conducted with the tablet application. The subject will first demonstrate a given task and take turns manipulating objects in the application scene with the robot (Figure 4). The task policy generated by the robot will be compared to the one that was originally intended by the human subject. The performance of the robot executing the learned task will be measured, and the subject will be asked to complete a survey that focuses on evaluating the interactiveness of the robot.

# 6. CONCLUSION

This thesis work focuses on developing an autonomous system that generates collaborative task-solving behavior through understanding a human's task. This work seeks to contribute to defining a general task-learning policy through observed visual task-learning cues. The system is applied to a play-task domain using a tablet application as a shared workspace between human and robot. Additionally, this work anticipates positioning the tablet workspace as a convenient multimodal-interaction research platform in the near future.

# 7. ACKNOWLEDGMENTS

My most sincere thanks go to my advisor and mentor, Dr. Ayanna Howard. I'd like to also thank my committee members Dr. Aaron Lanterman, Dr. Linda Wills, Dr. Andrea Thomaz, and Dr. Charlie Kemp for their comments on the proposal. This work was supported in part by the National Science Foundation under Award Number IIS-0705130.

# 8. REFERENCES

[1] G. Baranek, C. Barnett, E. Adams, N. Wolcott, L. Watson, and E. Crais. Object play in infants with autism: Methodological issues in retrospective video analysis. *American Journal of Occupational Therapy*, 59(1):20–30, 2005.

[2] C. Breazeal. *Designing sociable robots*. The MIT Press, 2004.

[3] D. Feil-Seifer and M. Mataric. Defining socially assistive robotics. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pages 465–468. IEEE, 2005.

[4] H. W. Park and A. Howard. Case-based reasoning for planning turn-taking strategy with a therapeutic robot playmate. In *Biomedical Robotics and Biomechatronics (BioRob), 2010 3rd IEEE RAS and EMBS International Conference on*, pages 40 –45, Sept. 2010.

[5] H. W. Park, A. Howard, and C. Kemp. Understanding a child's play for robot interaction by sequencing play primitives using hidden markov models. In *Robotics and Automation, 2010. ICRA '10. IEEE International Conference on*, May 2010.

[6] B. Robins, K. Dautenhahn, and P. Dickerson. From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot. *Proc. Second Inter. Conf. Advances in CHI, ACHI'09.*, 2009.

[7] T. Starner and A. Pentland. Visual recognition of american sign language using hidden Markov models. In *International Workshop on Automatic Face and Gesture Recognition*, pages 2982–2987, 1995.

[8] A. Thomaz, C. Breazeal, A. Barto, and R. Picard. Socially guided machine learning. *Computer Science Department Faculty Publication Series*, page 183, 2006.

[9] K. Wada, T. Shibata, T. Saito, K. Sakamoto, and K. Tanie. Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 2785–2790. IEEE, 2005.

[10] J. Wainer, D. Feil-Seifer, D. Shell, and M. Mataric. Embodiment and human-robot interaction: A task-based perspective. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pages 872–877. IEEE, 2007.