Learning Social Behaviors During Human-Robot Play

Cynthia Breazeal

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: This work explores socially situated learning by exploiting the types of interaction that arise between a nurturing caretaker and an immature learner. Here, the learner is an anthropomorphic robotic platform. Its primary sensory inputs include vision, audition, and joint rotation of its degrees of freedom. Its outputs include vocalizations, head and eye orientation, and facial expressions. The robot is designed to be an altricial system, similar in spirit to a human infant. That is, the robot starts off in a rather helpless and primitive condition, and requires the help of a sophisticated and benevolent caretaker to learn and develop. The interaction between the caretaker robot is purely social, much like how a mother interacts with her infant. The sorts of capabilities targeted for learning are those social and communication skills exhibited by human infants within the first year of life.

Motivation: This work is aimed at the larger question of how to build a more open ended learning system. Today, much of the learning-based research in robotics is targeted at training a robot to learn a specific task, model, representation, et cetera. Often the researcher decides a priori what task the robot is to learn (such as navigating around an office environment) and then sets out to engineer the learning task accordingly. The learning task is completed once the robot can perform the task to a desired measure of success. However, because the learning algorithm is carefully tailored to a specific task, a new learning algorithm must be painstakingly designed for the robot to learn a different task. The design of learning algorithms for robots is a labor intensive process, and it is proving difficult to scale current techniques to more complex tasks in more complex environments.

In contrast, this work explores how to design a more open ended learning system. To this end, it is heavily inspired by the theories, observations, and experimental results of child developmental psychology. The heart of this research is to figure out how to design an integrated learning system such that the learner can bootstrap from previously acquired skills and cognitive structures to learn new, more diverse, and more sophisticated skills. Human infants are the prime exhibitors of the kinds of learning we want our system to emulate, often characterized as having a developmental profile where earlier skills and competencies are progressively modified, adapted, and built upon to produce more sophisticated, diverse, or new abilities.

Approach: Our approach is designed to exploit many of the same social cues, constraints, and biases that human infants have access to as they learn to interact with a caregiver. Starting off in a relatively primitive state can be used to advantage when learning early skills, especially when the system is situated within an environment that is horrendously complicated. As with newborn infants, early learning problems are simplified as crude initial perceptual abilities and limited motor skills limit the complexity of the information that the system must deal with and learn from. This gives the infant a chance to learn skills that are commensurate with its current perceptual, motor, and cognitive abilities instead of bombarding the neonate with mind-numbing complexity. In addition, the caretaker helps simplify the learning problem for the infant by benevolently engineering the environment and the task to meet the infant's current capabilities. Examples include using brightly colored objects, drawing the infant's attention to salient factors, assisting the infant in doing the task at hand, and so on.

The interaction between learner and caretaker forms a mutually regulating process. Using emotional feedback from the infant, the caregiver orchestrates the learning episode to suit the learner's current level of sophistication. For instance, if the learner is over-stimulated (too overwhelmed by environmental complexity), the caretaker must simplify or even pause the learning episode. Alternatively, if the learner appears bored, the caretaker either introduces either a bit more variety or a little more difficulty to the learning episode. Over the course of learning, the learner constructs internal structures to implement more sophisticated skills and competencies. As a result, it is capable of handling a slightly more complex environment. Hence, a balance is maintained where the learner is always sufficiently challenged to learn beyond what it already has, but is never completely overwhelmed so that is has little chance of learning anything. As development proceeds, new goals are learned as interesting outcomes are discovered as well as different means to achieve them.

Difficulty: There are many difficult issues that must be addressed by this work. The two primary issues are establishing a natural, real-time dynamic of interaction between the learner and caretaker, and coming up with an appropriate set of learning mechanisms and substrate with which to build internal structures. The perceptual aspects alone are notoriously difficult problems, and we will have to trade-off perceptual sophistication for real-time performance in an effort to focus this research on the learning and interaction aspects of the problem (instead of becoming immersed in complex visual and auditory processing). What that trade off will be has yet to be determined.

Impact: From a broader perspective, this research not only aims at trying to build an open ended learning system, but also to build a system that humans can interact with and train in a natural, instinctive manner. Humans are highly social creatures and use a variety of cues and modalities to communicate with each other. Building systems that can exploit and understand similar social cues could make machines easier for people to use, and enable humans to communicate with machines in richer ways [2].

Future Work: This work is quite different from a robotics perspective in using inspiration from developmental psychology to learn communication skills in a social context. To date, a robotic platform (called Kismet) has been built. Kismet currently has an active stereo vision system with color CCD cameras, and a variety of facial features that allow it to display a wide range of facial expressions. It perceives visual motion, color, and the presence of faces, three important perceptual abilities of human newborns. An attentional system has been implemented which integrates this low level perceptual information with motivational and behavioral factors to direct the robot's gaze to the most salient stimuli [3]. The robot's motivational system models both drives and emotions, and the satiation of these drives and the maintainence of a positive emotional state motivate the learning process [1] and influence facial displays. Kismet's behavior system implements key proto-social responses exhibited by newborns to encourage the human cargiver to socially engage the robot. The robot's controller enables it to successfully engage a human in infant-like social interactions – a necessary precurser for socially situated learning.

Current work entails endowing Kismet with emotive vocalization capabilities and the ability to perceive the affective state of the caregiver through auditory input channels. This will provide Kismet with important reinforcement during social learning episodes. From there, most of the effort will explore the social developmental learning aspects of the problem.

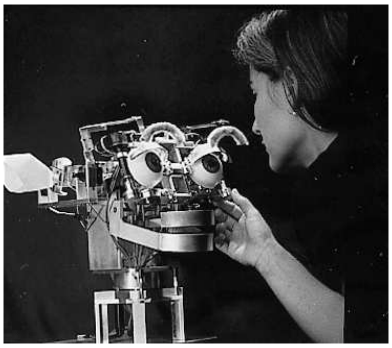

Figure 1: Kismet, an infant-like robot, has been designed to explore socially situated learning. The social context is similar to that of an infant-caregiver scenario, where the human caregiver shapes and guides what Kismet learns through social interactions.

Research Support: "A Trainable Modular Vision System" funded by ONR under contract number N00014-95-1-0600 administered by ONR.

References:

[1] C. Breazeal. A motivational system for regulating human-robot interaction. *Proceedings of AAAI98*, 1998, Madison, WI, pp. 54–61.

[2] C. Breazeal. Robot in society: friend or appliance? *Proceedings of Agents99 workshop on emotion-based architectures*, 1999, Seattle, WA, pp. 18–26.

[3] C. Breazeal and B. Scassellati. A context dependent attention system for a social robot. *Proceedings of IJCAI99*, 1999, Stockholm, Sweden, pp. 1146–1151.