Implementing Models of Autism with a Humanoid Robot

Brian Scassellati

The Problem: Autism is a pervasive developmental disorder that selectively impairs elements of social functioning [3]. One current model of autism points toward deficits of joint attention mechanisms, which allow individuals to respond to a variety of social cues such as pointing gestures, gaze direction, and postural cues that serve to focus attention of both individuals on a distal object [1]. The primary focus of this research is to investigate how individuals develop the necessary skills to recognize and produce joint attention by implementing models of the development of joint attention on a humanoid robot [2].

Motivation: One of the critical precursors to social learning in human development is the ability to selectively attend to an object of mutual interest. Joint attention to objects and events in the world serves as the initial mechanism for infants to share experiences with others and to negotiate shared meanings. Joint attention is also a mechanism for allowing infants to leverage the skills and knowledge of an adult caretaker in order to learn about their environment, in part by allowing the infant to manipulate the behavior of the caretaker and in part by providing a basis for more complex forms of social communication such as language and gestures.

Constructing mechanisms of joint attention for a robot provides both difficult engineering challenges and interesting scientific questions. From the engineering perspective, constructing a machine that can recognize the social cues from a human observer allows for more natural human-machine interaction, creates possibilities for machines to learn by directly observing untrained human instructors, and expands on the growing capabilities of robotic systems. Constructing a system capable of recognizing and exhibiting joint attention requires integrating components of many active areas of computer science research: machine vision (face recognition, gesture identification, etc.), visual-motor coordination, and machine learning. From a scientific perspective, an investigation of shared attention asks questions about the development and origins of the complex non-verbal communication skills that humans so easily master. Studies of child development, developmental disorders, and the evolution of communication all contribute interesting questions to this line of research: What is the progression of skills that humans must acquire to engage in joint attention? When something goes wrong in this development, as it seems to do in Autism, what problems can occur, and what hope do we have for correcting these problems? What parts of this complex interplay can be seen in other primates, and what can we learn about the basis of communication from these comparisons?

Previous Work: Joint attention has been investigated by researchers in a variety of fields. Experts in child development are interested in these skills as part of the normal developmental course that infants acquire extremely rapidly, and in a stereotyped sequence [5]. Additional work on the etiology and behavioral manifestations of pervasive developmental disorders such as autism and Asperger's syndrome have focused on disruptions to joint attention mechanisms and demonstrated how vital these skills are in human social interactions [3, 1]. Philosophers have been interested in joint attention both as an explanation for issues of contextual grounding and as a precursor to a theory of other minds [7]. Evolutionary psychologists and primatologists have focused on the evolution of these simple social skills throughout the animal kingdom as a means of evaluating both the presence of theory of mind and as a measure of social functioning [4].

Our research on this project has focused on the construction of sensory-motor skills (such as visuo-motor skills and hand-eye coordination), perceptual abilities (such as face and eye finding), and on the associated cognitive systems (such as motivational and attentional systems). This work is reviewed in [6].

Approach: Using a comprehensive model of the joint attention deficits in autism, we have developed a developmental decomposition of joint attention skills which will be implemented on our humanoid robot. These skills include recognition of mutual gaze, extrapolation of the angle of gaze to produce gaze alternation, imperative pointing (pointing to request an object) and declarative pointing (pointing to draw attention). This decomposition has support from evidence from child development [5], from studies of developmental disorders [1], and from ethological studies of social skills [4].

Difficulty: The primary difficulty in this work is integrating techniques from a wide variety of disciplines including machine vision (face recognition, gesture identification, etc.), visual-motor coordination, and machine

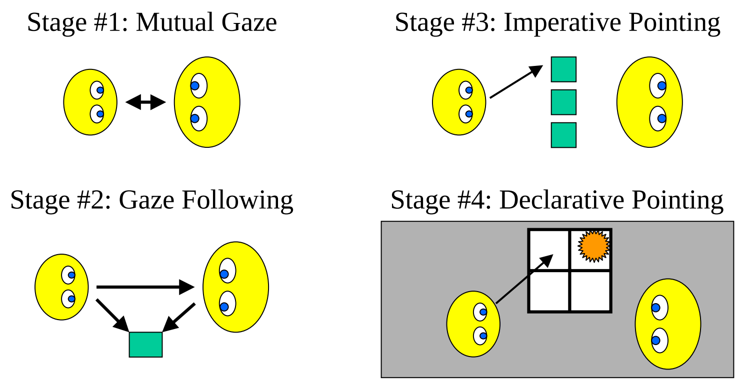

Figure 1: Elements of a developmental progression of joint attention skills.

learning.

Impact: Building a developmental model on our robot will give insight into our original biological observations. Implementing a model of joint attention will allow us to test and evaluate that model, uncovering inconsistencies and filling out unspecified details. With an implementation of joint attention, our robot can serve as a type of experimental apparatus that has been unavailable to the cognitive science community. Effects of different environments, different experiences, or even abnormalities of development can be investigated in isolation by simply changing the parameters of our robotic platform.

These mechanisms will also allow our robot to learn from people in a natural, unconstrained manner. Just as a child learns social skills and conventions through interactions with its parents, our robot will learn to interact with people using natural, social communication.

Future work: This research is just beginning. We currently have built perceptual systems for detecting faces, eyes, visual motion, and areas of high color saliency. We also have constructed motor control systems for a variety of eye movements (saccades, smooth pursuit, and VOR/OKN), neck movements (including orienting to salient objects with efference copy), and some preliminary arm motions (such as pointing to a visual target). We have also constructed an attentional system that integrates perceptual saliency with the robot's motivations and goals to direct limited computational and motor resources. Our next step is the detection of mutual gaze and finding gaze direction.

Research Support: Parts of this research are funded by DARPA/ITO under contract number DABT 63-99-1-0012 and parts have been funded by ONR under contract number N00014-95-1-0600, "A Trainable Modular Vision System."

References:

[1] S. Baron-Cohen. *Mindblindness* MIT Press, 1995.

[2] R. A. Brooks, C. Breazeal (Ferrell), R. Irie, C. C. Kemp, M. Marjanovic, B. Scassellati, M. Williamson. Alternate Essences of Intelligence. AAAI 98.

[3] D. J. Cohen and F. R. Volkmar, eds. *Handbook of Autism and Pervasive Developmental Disorders, second edn.* John Wiley and Sons, Inc., 1997.

[4] D. J. Povinelli and T. M. Preuss. Theory of mind: evolutionary history of a cognitive specialization. *Trends in Neuroscience*, 18:9:418–424, 1995.

[5] M. Scaife and J. Bruner. The capacity for joint visual attention in the infant. *Nature* **253**, 265–266, 1975.

[6] B. Scassellati. Investigating Models of Social Development using a Humanoid Robot. AAAI Fall Symposium, 1998.

[7] A. Whiten, ed. *Natural Theories of Mind.* Blackwell, 1991.