Instructing an attentive robot

Paul Fitzpatrick

The Problem: This work looks at teaching an anthropomorphic robot simple sequential, looping, and branching processes. It will concentrate on the use of instruction rather than relying on autonomous learning.

Motivation: A robot in the presence of an instructor who can monitor and influence its attention can be taught a great deal more than it could learn independently. Visual attention channels a robot's experience of the world along a linear sequence. This constraint provides an instructor with useful leverage for directing the robot's learning – for example, by manipulating the saliency of objects or by using mechanisms of joint attention. But equally importantly, the existence of this linear sequence makes it easier for an instructor to follow the activity of the robot's learning mechanism. In an anthropomorphic robot with foveated vision, the physical locus of attention of the robot can be determined from its gaze direction. Hence the instructor can track what the robot's learning mechanism is truly being presented with – as opposed to what the instructor intended to teach. Hence, if the robot indicates confusion, the instructor is not at a loss to diagnose the source of the confusion and modify the teaching strategy accordingly.

Previous Work: Scassellati has developed some initial steps towards joint attention [3]. Breazeal and Scassellati have developed an attentional system that is sensitive to people and toys [4], and are also working on an imitation system. A number of basic eye movements have been implemented for Cog, including vergence, saccades, smooth pursuit, and the vestibulo-ocular reflex. Work remains to be done developing robust gaze where the different types of eye movements are tightly integrated. Other implementations of gaze systems include [1] and [5].

Approach: The learning mechanism for this work operates more as a means of communication than as a form of machine learning in the conventional sense. Hence it is important that the learning mechanism be incremental, and easily predicted by a human instructor. One choice that appears to lead to this is to model processes as being generated by a hierarchy of non-deterministic finite automata that successively refine branching choices.

The robot will not be expected to perform the processes it is taught, but rather to predict them. For example, it may be asked to watch a series of actions and predict what will happen next, based on previous examples. In particular, variants on the "shell game" will be used, where the robot tries to track a hidden ball.

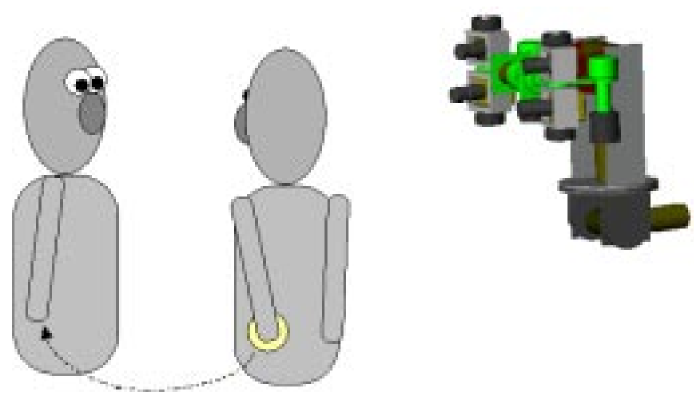

Figure 1: The human shell game – participants pass or dummy-pass a ball to each other behind their backs. The robot watches, and looks at the person it believes has the ball.

Difficulty: There are many methods for learning finite automata, but the challenge is to find one that will give results close to the instructor's intuitive expectations of what should be learnable from the examples she presents. This work relies on a number of technologies co-operating together – mechanisms of joint attention, visual tracking, gesture recognition etc. – none of which are themselves mature.

Impact: This work is a first step towards enabling Cog to develop expectations about the world, and how events sequence and relate.

Future Work: This is just beginning. The processes this work is concerned with are very limited, particularly in the kinds of dependencies that may exist between actions. This would need to be improved before groups of processes could be combined in a robust manner. But in practice most effort is likely to be directed at resolving perceptual difficulties, such as how to perceive a process in the same way as the instructor.

Research Support: This work is funded by DARPA/ITO under contract DABT 63-99-1-0012.

References:

[1] A. Bernardino, J. Santos-Victor. Visual behaviours for binocular tracking. Robotics and Autonomous Systems 25, 1998.

[2] Rodney A. Brooks, Cynthia Breazeal (Ferrell), Robert Irie, Charles C. Kemp, Matthew Marjanovic, Brian Scassellati, Matthew Williamson. Alternate essences of intelligence. AAAI 98.

[3] B. Scassellati. Imitation and Mechanisms of Joint Attention: A Developmental Structure for Building Social Skills on a Humanoid Robot. *Computation for Metaphors, Analogy and Agents*, Springer Lecture Notes in Artificial Intelligence, Vol. 1562, 1999.

[4] C. Breazeal and B. Scassellati. A Context-Dependent Attention System for a Social Robot. IJCAI, 1999.

[5] S. Rougeaux and Y. Kuniyoshi. Robust Real-Time Tracking on an Active Vision Head. IROS 97.