Early Pragmatic Language Development for an Infant Robot

Paulina Varchavskaia

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: How can a developing creature acquire natural language in a meaningful way? At an early stage of a human infant's life, certain verbal behaviors - both instinctive and learned - enable the infant to begin acquiring the caretakers' system of communication. The problem is to provide an infant robot with a mechanism for exhibiting such verbal behaviors, which would also have the function of acquiring words of a natural language.

Motivation: Motivation comes from the general direction of research on Kismet[1], a robot built to be an embodied developing agent situated in the human world, whose behavior follows its own internal agenda and reacts to the world and human presence in it. Like a human infant, Kismet was designed to elicit nurturing responses from humans, which would scaffold its development and learning. As the robot acquires more of the basics of a natural language, communication with humans becomes easier, and the range of possible concepts and tasks to learn increases. The approach taken here, based on the pragmatics of language acquisition, is motivated by a desire to produce a system that acquires words together with functional meanings.

Previous Work: We draw extensively on research in the fields of developmental psychology, phonetics and linguistics, and more recently, behavioral robotics and ethology. Some researchers [6, 7] approached the problem of acquisition of natural categories and labels by robots from the point of view of perceptual grounding. The robot analyzes the visual scene and the speech stream into segments, the best correlation between which will form a perceptual concept-label pair which is acquired by the robot. The project also utilizes the concepts of proto-linguistic behaviors [4] and referential mapping [5].

Approach: The emphasis of this project is on the pragmatic functionality of acquired concepts. The bases of the natural language that Kismet is developing are grounded not only in its sensory streams but also, primarily, in its "drive" to manipulate its environment in such a way as to achieve the satiation of its "desires". After Halliday [4], we identify a number of requisite behaviors in early linguistic development: request, calling, greeting, self-assertion, question, naming, and categorization. The development starts with a finite set of pre-programmed vocal responses to certain "emotional" states such as frustration. They are activated as responses to combinations of releasers, which signal the presence of visual and audio stimuli as well as internal state such as the currently active behavior. These proto-verbal responsive behaviors are arranged in an architecture inspired by the subsumption architecture [2] with the additions of randomness, state in the form of short-term and long-term memories, and timings (Fig.1), and are implemented using a script language Zac [3] designed for programming robotic control architectures. These responses establish an initial regulatory system of communication between the robot and the caretaker.

The next step up from such a closed formulaic and manipulative system is the actual acquisition of adult-like words and their appropriate use in the scenarios of naming and categorization. This stage of development is achieved through a reinforcement learning mechanism operating on a network of concepts. Nodes can be dynamically created and deleted and this updating of the network constitutes the learning in the system. Nodes are activated by recognizing the presence of specific releasers to which they are sensitive. Releasers here can be thought of as virtual sensors with particular receptive fields, or filters, which may be hierarchically organized.

Many nodes may be more or less active at any one time. What Kismet actually says is determined in a distributed way through the timings of node activations.

The robot learns the language on different planes. New reinforcement signals (e.g., labels such as "good" and "bad") are learned through correlation with pre-programmed ones. New meanings are created from combinations of releasers and internal virtual sensors. Labels for these are learned by a reinforcement algorithm. These meanings

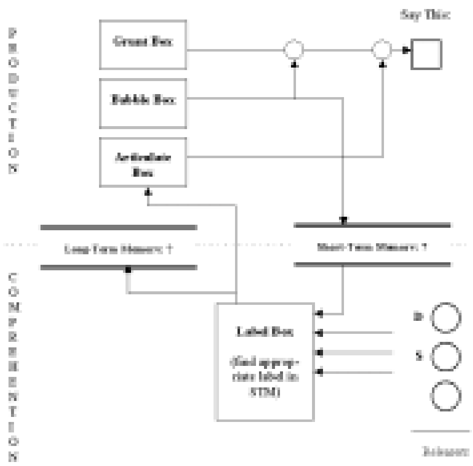

Figure 1: Proto-verbal behaviors on Kismet

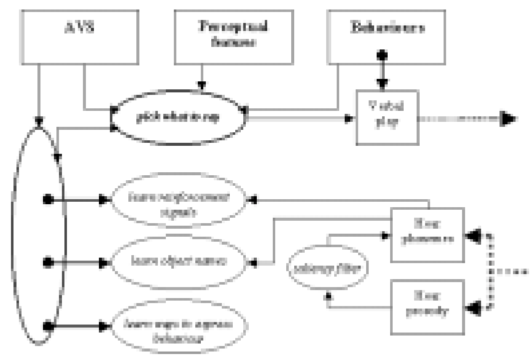

Figure 2: Acquiring new early concepts

are represented and expressed as new verbal behaviors, which constitute the pragmatics of a natural language.

Impact: Spoken language is one of the most powerful social tools for humans. When transferred to the domain of human-robot interaction, it has the advantage of requiring no special skills from the caretaker, who intuitively interprets the communicative acts as meaningful. Potentially, this approach can be pursued in developing social interfaces for end-user oriented applications and artificial creatures.

A robot that learns the language in the same way as children do will not be manipulating meaningless symbols but using its linguistic knowledge to express its needs and manipulate the world. An observer could describe such language at all levels, from pragmatic intentions of the speaker to the phonemes that are produced and heard. The whole exercise will pave the way for fully socially able machines.

Future Work: Numerous extensions are always possible in the field of natural language acquisition by machines. In particular, this work does not address issues of more complex conceptual structures, or syntax. These issues should be examined in the future.

Research Support: This research is supported in part by DARPA under the “Natural Tasking of Robots Based on Human Interaction Cues” project contract number DABT 63-00-C-10102, and in part by the Nippon Telegraph and Telephone Corporation under the NTT/MIT Collaboration Agreement.

References:

[1] Cynthia Breazeal. *Sociable Machines: Expressive Social Exchange Between Humans and Robots*. PhD thesis, MIT, May 2000.

[2] Rodney Brooks. A robust layered control system for a mobile robot. A.I. Memo 864, MIT, 1985.

[3] Paul Fitzpatrick. A novel behaviour-based robot architecture and its application to building an autonomous robot sentry. Master’s thesis, University of Limerick, Ireland, 1997.

[4] Michael Halliday. *Learning How To Mean: Explorations in the Development of Language*. Cambridge: MIT Press, 1975.

[5] Irene Pepperberg. Referential mapping: A technique for attaching functional significance to the innovative utterances of an african grey parrot. *Applied Psycholinguistics*, (11):23–44, 1990.

[6] Deb Kumar Roy. *Learning words from Sights and Sounds: A Computational Model*. PhD thesis, MIT, September 1999.

[7] P. Cohen T. Oates, Z. Eyeler-Walker. Toward natural language interfaces for robotic agents: Grounding linguistic meaning in sensors. *Proceedings of the 4th International Conference on Autonomous Agents*, pages 227–228, 2000.