BBS Supplemental Commentary

Below are links to the 15 supplemental commentaries on the target article, "From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics" by Michael A. Arbib.

Author's Electronic Response to the 15 Supplemental Commentaries: Supplemental Electronic Response

---

Breitenstein

Meaning as social norm
Martin Ruchsow

The "complex first" paradox: What's needed to lexicalize neurally distributed meanings?
Markus Werning

Assisted imitation: Linking action sequences, communication, and the emergence of language
Patricia Zukow-Goldring

Commentary on Arbib

Word Counts:
Abstract: *55* words
Main Text: *973* words
References: *78* words
Total Text: *1106* words

A Deeper Semantic Role for the Mirror System

Yoonsuck Choe
Department of Computer Science
Texas A&M University
3112 TAMU
College Station, TX 77843-3112
U.S.A.
+1-979-845-5466
choe@tamu.edu
http://faculty.cs.tamu.edu/choe

Abstract: In proceeding from simple communication to full-blown language, it may have been critical to better understand one's own internal state, as much as it was to infer that of others. The mirror system may have allowed such a more elaborate understanding of oneself just like it allowed the understanding of more complex gestures/utterances of others.

In the target article, Arbib advances two important points regarding the mirror system. First, the mirror system supports the parity property (Section 2.1), and second, such a property can be learned through experience (Section 3.2). The parity property basically states that ''what counts for the speaker (or producer) must count for the listener (or receiver).'' Thus, its concern is mostly about inferring the intention of others. An interesting question is, can this property also apply to the first-person, i.e., can the parity property be applied in understanding one's own internal state (such as a perceptual state)? Another question is, if such a property is to be learned, what kind of criterion (other than minimizing the obvious input-output mapping error) can be used to guide the learning process?

A simple thought experiment allows us to think more clearly about both of these issue (Choe and Bhamidipati 2004). For example, consider a simple perceptual agent illustrated in Figure 1. The agent has a single visual filter *f*, which is a function *f:I* *S*, where *I* is the current input in the visual field and *S* is the output state value (e.g., a neuronal spike which can be either 0 or 1). Suppose this visual filter *f* is highly tuned to an oriented line at 45° (i.e., its receptive field has a 45° orientation preference). Given only this, can this agent figure out that *S* was triggered by an external perceptual input that is a 45° line, just by looking at *S*, and that without any knowledge about the particular form of the function *f*? In other words, can the agent figure out the meaning of its internal state *S*? Obviously, given only *S*, such an agent will not be able to infer the property of *I* because *S* is just a spike (0 or 1), an arbitrary ''sign''. For example, someone inside the box in Figure 1 cannot even answer an easier question of whether *S* was caused by visual input or auditory input (or for that matter, any other kind of sensory input; cf. Sharma et al. 2000). A clue can be found in the sensory substitution experiments by Bach-y-Rita (1972), where a tactile display was perceived to be visual in nature by a blind subject when voluntary motion was permitted, to dynamically alter the perceived tactile state coupled to a video camera.

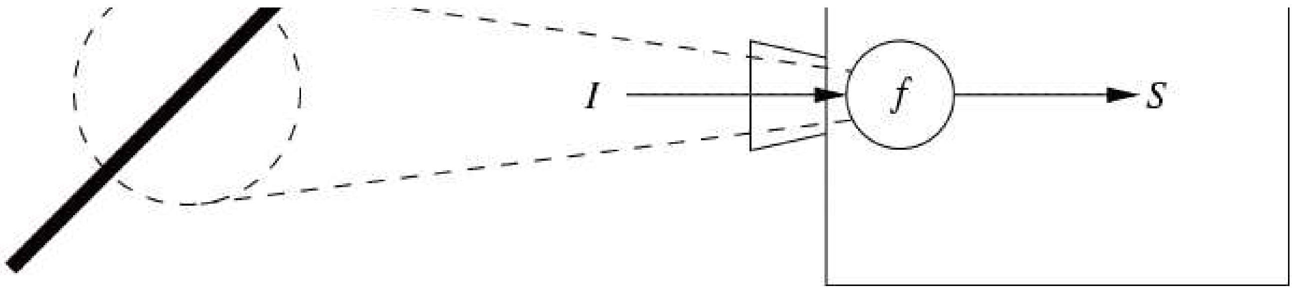

Figure 1. A simple perceptual agent

Thus, the only solution to this problem may be to allow the agent to generate action, e.g., gazing at different locations of the input by actively moving around its visual field. The key insight here is that the kind of motion that keeps the state *S invariant* over time during the execution of action exactly reflects the oriented property of the input stimulus *I* (and hence that of the filter *f*). For example, in Figure 1, the only motion that can keep the state *S* invariant (i.e., fixed at 1) is to move the visual field back and forth in the 45° diagonal, and such a motion exactly reflects the stimulus property encoded by *f*. Thus, even without any knowledge about *I* nor about *f*, the agent can learn (in an *unsupervised* manner) what its internal state *S* means through action that maintains invariance in *S*. Based on the above, two points become clear: (1) voluntary action can provide meaning to one's internal perceptual state, and (2) maintained invariance in the internal perceptual state can serve as a criterion (or an objective) for learning the appropriate action sequence. It is easy to see how the example discussed here can be extended to gesture recognition (the function *f* for the gesture *I*) and the generation of an analogous act (action based only on *S*), where invariance in the mirror neuron may serve as a learning criterion. Thus, it may be the case that mirror-system-like perception-action coupling may be playing an important semantic, epistemic role about one's own internal state. In other words, the parity property can be applied internally to better understand one's own internal state: what counts for my voluntary actions counts for my own perceptual states. An interesting question is whether it is also possible that such a development preceded that of communication and language: one must possess (raw) semantics before any form of communication (with others) becomes possible. If this is true, Arbib's rejection of the ''mirror system for concepts'' (Hurford 2003a) in Section 5 may have to be reevaluated, because the mirror system may in fact be involved in the neural representation of the ''signified,'' and not just the ''sign.'' Arbib may be right in that ''mirror schema'' only occurs very rarely, but we can also think of these mirror schemas as forming the semantic building block for a more open ended ''schema network of concepts.'' Thus, beyond the mirror system, a population of neurons (not just a single neuron) may take on the mirror property as it becomes necessary to grasp more complex, combinatorial concepts, and such a population may dynamically change over time as needed.

In sum, the mirror system can be used not only for the communication between individuals, but also for a better understanding of one's own internal state through action if its perception-action coupling function is directed inward. Such a form of self-understanding may be present in animals at all levels, but it may have become more complex and elaborate in humans as the mirror system progressed from imitating a simple to a complex set of actions (Section 1.2, S3 to S4). I.e., the depth of self-understanding in humans may have grown exponentially in a ''spiral,'' tangled together with the emergence of language. Finally, the discussion above also sheds some light on how the mirror property can be learned through experience, and what kind of learning criterion may have been used (i.e., that of sensory invariance).

References

Bach-y-Rita, P. (1972) *Brain Mechanisms in Sensory Substitution*. Academic Press, New York, NY.

Choe, Y. and Bhamidipati, S. K. (2004) Autonomous Acquisition of the Meaning of Sensory States Through Sensory-Invariance Driven Action. In: *Biologically Inspired Approaches to Advanced Information Technology* eds. Ijspeert, A. J. and Murata, M. Springer, New York, NY.

Hurford (2003a): See the reference in the target article.

Sharma, J., Angelucci, A., and Sur, A. (2000) Induction of Visual Orientation Modules in Auditory Cortex. *Nature* 404:841—847.