Learning the Meaning of Neural Spikes Through Sensory-Invariance Driven Action

Yoonsuck Choe and S. Kumar Bhamidipati
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
{choe,bskumar}@tamu.edu

Abstract

What does a spike of a sensory neuron mean? This is a fundamental question in neuroscience, and usually it is answered by relating neural firing with the stimulus that triggered the spike. However, the problem becomes difficult if we pose this question to another neuron inhabiting the same brain, because that neuron does not have immediate access to environmental stimuli. We reduce this problem into a small task, and propose that *action* is necessary for acquiring the *meaning* of neural spikes. Similar approaches exist, but our approach differ in a novel way that the objective of the agent is to learn action that maintains *sensory invariance*. We implemented this idea in a reinforcement learning agent, and the results show that the agent successfully learns to associate sensory neuron activity to action resembling the stimulus that maximally activates that neuron. Thus, the implication is that the meaning of a sensory neuron spike can be found in the pattern of action that reliably activates that neuron. This approach is not limited to a single neuron, as it can be easily extended into spatial and temporal patterns of firing. We firmly believe that further investigation into this direction can be most fruitful.

1 Introduction

The brain is made up of 100 billion neurons, which generates a complex pattern of activity in response to sensory stimuli from the external environment. A fundamental question in neuroscience is, how do we understand what this pattern means? To make the question even simpler, we can ask what does a spike of a single neuron *mean*? [1]. Even this reduced problem is not trivial, and it took an enormous effort to come to the current state of understanding, beginning from muscle spindles [2] to cat visual cortical neurons [3] to sophisticated stimulus reconstruction methods developed lately (see, e.g., [1]).

A popular approach to this question is through associating the neural spikes with the stimulus that triggered those spikes [4] (see [5] for a review). Such methods have been successful in characterizing the neural spiking properties and accurately predicting the stimulus given just the spike train. This method involves the experimenter systematically varying the environmental stimulus while measuring the neural response (see, e.g., [1] chapter 2), so that at a later time, when only the spike train is observed, something can be said about the stimulus

property. Mathematically, this is conveniently written using the Bayes theorem [1]:

$$P(s|t) = \frac{P(t|s)P(s)}{P(t)},$$

where s is the stimulus and t is the spike train. Note that the likelihood term $P(t|s)$ requires that we have either an empirical statistics or a reasonable model of the stimulus-to-spike translation. Thus, the interpretation of the current spike train $P(s|t)$ seems to *depend* on knowledge about the stimulus properties (from the past).

Now suppose we ask the same question "what does a single spike mean?" to *another* neuron in the brain where such spikes are received. Because this neuron does not have immediate knowledge about the environmental stimulus associated with the spike it received (as it does not have the tools of an experimenter) the neuron cannot apply the technique mentioned above. (This problem can also be seen in the context of the Bayesian theorist, i.e., not merely an observer; an issue raised by Jepson and Feldman [6].) For example, consider a similar situation depicted in figure 1. Inside the agent, the only available information is the sensory array activity, so if we are trapped inside this agent without access to the visual environment outside, we can never figure out what the sensor activity means.

This paper begins by realizing this as a genuine issue. Our contention is that such a dilemma can be overcome through learning how to associate sensory activity to the motor actions the brain itself generates. The importance of sensory-motor learning has been emphasized in the past by many researchers, for example, (1) sensory-motor contingency theory [7], (2) two-level interdependence theory [8], (3) ecological perception [9], (4) subsumption architecture in robotics [10], (5) learning of natural semantics [11], and (6) mirror neurons and imitation [12, 13]. However, to our knowledge, the question of attaching *meaning* to a single spike has not been addressed, nor has a principled objective for learning been proposed in this particular context.

Below, we first define the problem in terms of the agent we introduced in figure 1, and propose a learning algorithm based on *sensory-invariance driven* motor action. The basic idea is that the agent has knowledge about its own movements, and the movements that it generates that reliably activate a particular sensor in the sensor array constitute the meaning of that sensor's spike. The acquired meaning for each sensor and the resulting behavioral patterns are presented next, followed by discussion and conclusion.

2 Meaning of Spikes in a Sensory-Motor Agent

To better illustrate our point, let us consider a small, concrete example as shown in figure 1, a simple sensory-motor agent. The agent has a limited visual field, and the incoming visual signal is transformed via the oriented filters (mimicking primary visual cortical neurons) into a spike pattern in the *sensory array*. Let us further assume that the agent does not have any knowledge (e.g., about the receptive field structure) of its oriented filters. The task of the agent then is to attach meaning to its own sensory array activity pattern, i.e., to come to the understanding that each sensor represents a certain oriented visual input.

Imagine we are inside this agent, isolated from the world outside the box, sitting near the big "?" sign. It is questionable then whether we can ever be able to associate an orientated visual input stimulus with the spikes generated in the sensor array because we cannot peek outside, and we do not know the particular mechanism of the filters. The spike, in principle, could have been generated from any sensory modality, e.g., auditory or tactile input.

The only way we can see this issue resolved is through action, that is, the movement generated by the agent. This point is not entirely obvious at first, so let us elaborate a little bit what we mean. As shown in figure 1, we included the capability of action in the agent. The
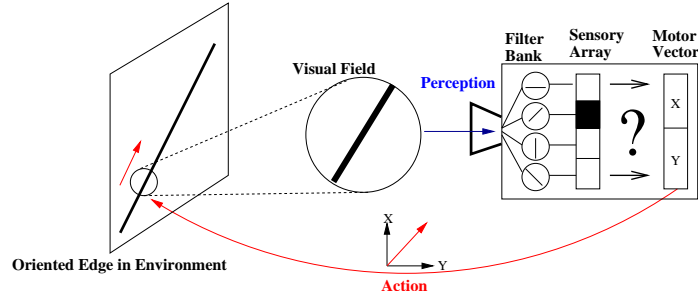
Figure 1: **A Sensory-Motor Agent.** An illustration of a simple sensory-motor agent is shown. The agent has a limited visual field where the input from the environment is projected. A set of orientation-tuned neurons receive that input and generate a pattern of activity in the sensory array (black marks active). In the situation here, the $45°$ sensor is turned on by the input. Based on the sensory array pattern, after some processing (signified by "**?**"), the x and y values of the motor vector is set, resulting in the movement of the visual field and a new input is projected to the agent.

agent is able to gaze at different parts of the scene by moving around its visual field. The x and y variables correspond to the movement of the visual field in the x and the y direction, respectively. Thus, these two variables are like motor commands. We, sitting on that "?" sign, can generate different combinations of (x, y) values and observe the changing pattern in the sensory array. By relating the sensor activity and the motor command that was just generated, *certain aspects* of the sensor property can be recovered. We believe this is generally agreeable, but it is too general. It begs the question of *what* is that "certain aspects" of the sensory property and *how* can they be learned?

A crucial insight that occurred to us at this point was that certain kinds of action tend to keep the sensory activity pattern to remain unchanged (i.e., *invariant*) during vigorous movement, and this action exactly reflects the property of the sensory stimulus. For example, consider the state of the agent as shown in figure 1, where a $45°$ input is presented, and the corresponding sensor is activated in the agent. Now imagine we move the visual field according to the motor vectors $(1, 1), (1, 1), ..., (1, 1), (-1, -1), (-1, -1), ..., (-1, -1)$, which corresponds to a back-and-forth movement along the $45°$ diagonal (i.e., aligned on the input). Such an action will keep only the $45°$ sensor turned on during the motor act, i.e., the sensory array stays invariant. We can see that this motion, generated while trying to keep the sensor array unchanged, has led the agent to perform an act which reflects the stimulus. Thus, we are led to conclude that associating this kind of sensory-invariance driven action with spikes can potentially serve as a *meaning* for each sensory neuron.

To test this insight that ascribing meaning to sensory neuron activity is possible through learning the sensory-motor association based on sensory-invariance, we implemented a learning agent following the description in figure 1. The following sections describe the learning rule of the agent, followed by the results.

3   Learning of Sensory-Invariance Driven Action

Consider the agent described above (figure 1). We define a simple learning rule based on our idea of sensory-invariance driven action. The agent has the current state of its sensors S (the sensory array), and a set of actions D (possible combinations of the motor vector) that it can perform. For simplicity, we limit the sensor state set S to four different values

$$S \equiv \{0°, 45°, 90°, 135°\}, \tag{1}$$

which correspond to the four different orientation preference (note that $0°$ is the same as $180°$ etc.) of the sensors, and the action set D to eight different categories

$$D \equiv \{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}, \tag{2}$$

which are the possible *directions* of motion of the visual field with a limited distance of movement. Thus, the above corresponds to $(1, 0), (1, 1), (0, 1), (-1, 1), ...$ in terms of the motor vectors we mentioned earlier (the motion in x and y were either 0 or 1).

The learning task of the agent can then be treated as a standard reinforcement learning problem with a Markov assumption (see, e.g., [14]). The goal of the agent is to select an action from the action set D that maintains the sensory array activity *invariant*. Thus, the reward is simply the degree of sensory-invariance in successive stages of action. More formally, the agent has to learn a policy function π,

$$\pi : S \rightarrow D, \tag{3}$$

at step t which selects a direction of motion $d_t \in D$ based on the previous state $s_t \in S$ so that the resulting reward r_t is *maximized*. The execution of the policy at each state s_t results in reward:

$$r_t = r(s_t, d_t), \tag{4}$$

based on the reward function $r(s, d)$ for $s \in S, d \in D$, and this function is updated as follows:

$$r_{t+1}(s, d) = \begin{cases} r_t(s, d) + \alpha * f_t & \text{if } s_t = s_{t-1}, \\ r_t(s, d) - \alpha * f_t & \text{if } s_t \neq s_{t-1}, \end{cases} \tag{5}$$

where r_{t+1} is the reward at step $t + 1$; $\alpha (= 0.01)$ is a fixed learning rate; and f_t is the number of action steps taken by the agent up till t which resulted in either (1) continuously maintaining the sensory array to be invariant or (2) the opposite (i.e., changing all the time). Thus, if $s_t = s_{t-1}$ was true for the past n (= a large number) consecutive steps, then $f_t = n$, and this will increase the reward associated with (s, d). On the other hand, n consecutive failures of maintaining sensory invariance will also lead to a high f_t value, but this time the reward for (s, d) will decrease. The reward function is simple but even such a simple rule is sufficient for the agent to learn sensory-motor associations.

In the following, we will present the learned policy π and the behavior of the agent which mimics the input stimulus.
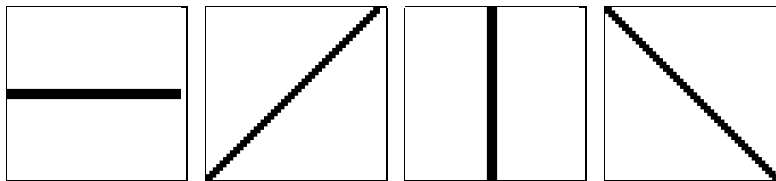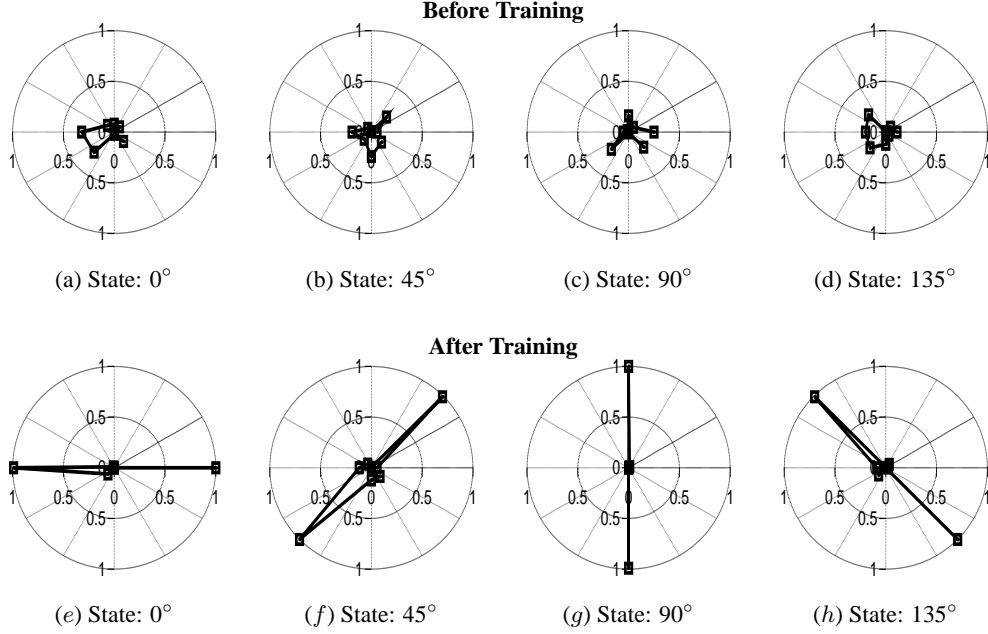
4 Experiments and Results

Figure 2: **Inputs Used for Training and Testing.** The agent was trained and tested on 51×51 bitmap images each containing a 3-pixel wide oriented edge. Four inputs with four different orientations are used for the experiments (from the left: $0°, 45°, 90°,$ and $135°$).

In the learning process the agent interacted continuously with the visual environment in a series of episodes. During each episode, the agent was presented with a 51×51 bitmap image containing an oriented edge (figure 2). The visual field of the agent was 9×9 which can slide across the image. The visual field input was directly compared to each of the four
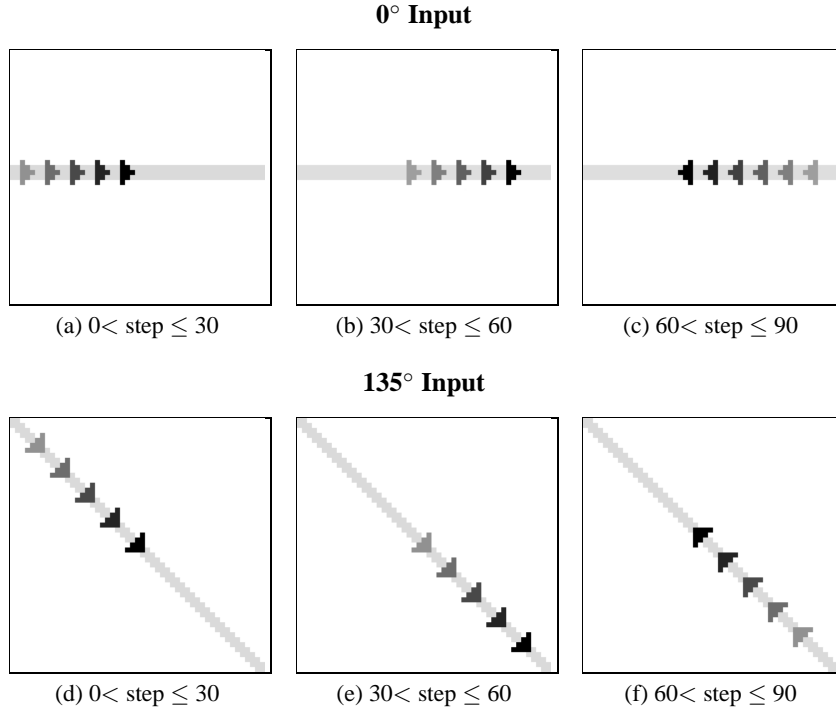
Before Training

(a) State: $0°$ (b) State: $45°$ (c) State: $90°$ (d) State: $135°$

After Training

(e) State: $0°$ (f) State: $45°$ (g) State: $90°$ (h) State: $135°$

Figure 3: **Reward Vector of Each Sensory State.** The reward values of the four possible sensory states ($0°$, $45°$, $90°$, and $135°$) are shown in polar coordinates. The top row from (a) to (d) are before training, and the bottom row from (e) to (h) are reward values after training. In each plot, for each point (θ, δ), the angle θ represents the direction $d \in D$ of the visual field movement (there are 8 possible directions), and the distance δ from the origin represents the associated reward value given the current sensory state (shown below each plot). The reward values were between 0 and 1. Initially, the rewards are randomly assigned for each direction of motion for each sensory state. After the agent is trained, the reward values become maximal for the movement along the orientations that correspond to the input that trigerred that sensory state.

sensory filters (also 9×9 in size) and the sensory state s was set to a value θ when there was an exact match with one of the four orientations $\theta \in S$ (see equation 1) .

The agent was trained to learn the policy $\pi : S \rightarrow D$ using equation 5 by going through the four different inputs. Since the size of the state and the action sets were $|S| = 4$ and $|D| = 8$, the policy π and the associated rewards can be enumerated in a 4×8 table. At each step, the next direction of motion $d \in D$ (see equation 2) was determined based on the expected reward values stored in such a reward table of the agent. The reward table was initialized to hold uniformly distributed random numbers between 0 and 1. Note that the reward was limited to the range $0 \leq r_t \leq 1$. Figure $3a$–d shows the initial reward values where each plot corresponds to a state $s \in S$, and each polar plot shows the reward r (distance from origin) for each action $d \in D$ (angle) for the given state s.

The training was carried out until the agent was able to learn to maximize the reward by consistently meeting the sensory-invariance criterion. The training usually lasted for up to 500 steps for each input. The reward table after training is visualized in figure $3e$–h. The results clearly show that the agent learned to associate motion d which reflects (or mimics) the actual orientation of the environmental input to the current sensory state s triggered by that input. For example, in figure $3f$, the maximum reward values associated with the sensory state $s = 45°$ are $d = 45°$ and $d = 225°$, indicating a preference for a back-and-forth movement along the $45°$ axis which exactly mimics the visual input. The same is true for all other states (figure $3e$, g, and h).

One thing to note from the actual numerical reward values (not shown) is that there is a

0° Input

| (a) 0< step ≤ 30 | (b) 30< step ≤ 60 | (c) 60< step ≤ 90 |

135° Input

| (d) 0< step ≤ 30 | (e) 30< step ≤ 60 | (f) 60< step ≤ 90 |

Figure 4: **Behavior of the Agent after Training.** Each plot shows a snapshot of 30 steps of movement of the agent's visual field in the 51×51 scene (only every 6 steps are shown). The triangles indicate the location of the visual field in the scene and their grayscale values represent the simulation step (black is the most recent step). The light gray lines in the background show the oriented input edges. Two simulation runs are shown here: (a) to (c) are for $0°$ input and (d) to (f) are for $135°$. The trained agent successfully generates motion sequence to trace the input in both runs based on its sensor state and policy π. For example, in (b) the agent starts in the center and moves right, and bounces back when it reaches the end of the input (c).

slight difference (≤ 0.01) between reward values for the two opposite directions separated by $180°$ (e.g., $d = 45°$ and $d = 225°$). The minor difference helps the agent to have an initial bias in the selection of the first movement, and to maintain a momentum to continuously follow along an orientation instead of rapidly oscillating between two opposite directions. Note that this desirable effect was not explicitly built in by us, but rather, emerged from the sensory-invariance driven learning rule.

In order to verify if our analysis of the reward table is accurate, the trained agent was tested with fixed oriented inputs and the resulting motor behavior was observed. Figure 4 shows the action sequence generated by the agent for two different inputs with orientations $0°$ and $135°$. The plots show the movement of the visual field of the agent in response to the given input. The results show that the action of the agent based on the learned reward table exactly reflects our analysis above: The agent, upon activation of a single orientation sensor, performs a movement mimicking the external input that triggered that sensor, thus assigning (in our interpretation) a *meaning* to the sensory neuron's spike in terms of its own action.

5 Discussion and Future Work

The main contribution of our work is the realization that a sensory-motor agent can find the meaning of its neural spikes within its own actions, and that the objective of maintaining sensory-invariance plays a key role in allowing the agent to autonomously discover this

semantic link.

An important message implicit in our work is that invariance can be seen from a totally different perspective. Usually, invariance is seen as something that needs to be detected by the perceptual system (e.g., invariant feature detection in vision). However, our approach differs in that invariance is seeked after in the neural activity pattern and it is *enforced* through a well-choreographed action. We speculate that there may be a link between this kind of action-based neural invariance and invariant sensory features in the conventional sense. For example, an approaching object will expand as time flows (turning on a certain neuron), and the same kind of effect can be achieved through a forward motion (again turning on the same neuron). Thus, the meaning of that neuron firing can be understood in terms of the action that would turn on that neuron reliably (cf. Gibson's work on ecological perception [9]). Thus, even without action, when that neuron turns on (i.e., object is approaching), the brain can infer that it is analogous to moving forward towards the object.

Bell [15] posed an interesting question regarding the perception-action cycle. To quote, "What quantity should a perception-action cycle system maximize, as a feed-forward channel might maximize its capacity?", which is relevant in our context. This is an important question, and we believe our reward criterion of *maximizing sensory invariance* can serve as a potential answer. As we have seen, such a criterion can be used to internally learn the meaning of sensory neuron firing which may be a very important function for a "perception-action cycle system" to possess.

One criticism we anticipate is that if the agent had a rich array of sensors, such as a 2D matrix of RGB pixels, then the properties of the visual environment can be easily recovered within the agent through unsupervised learning even without direct access to the outside world. However, this does not help solve the problem, because this rich information is only available at the very first stage of sensory processing. The next stage, and the stage following that, etc. only receive a more and more *encoded* version from the previous stage, just like the sensory array in our agent which receives only encoded spikes from the orientation-tuned filters. Thus, the same difficulty can remain.

A limitation of our account is that our model implicitly assumes that the agent has direct knowledge about its own movement, upon which the meaning of the sensors are grounded. We are not sure how this issue can be resolved, but the work by Philipona et al. [16] points into a direction where a possible resolution can be found. They showed that without any knowledge of the external world, physical properties of the environment can be learned through sensory-motor learning. This involves the understanding of its own actions, and here we hope to find a solution to our dilemma.

Can our approach be extended into other sensory modalities such as audition, somatic sense, olfaction, etc.? Our approach is general enough to be easily extended into certain modalities such as somatic sense, but it cannot work very well in domains where there is not much correlation between action and the perceived sensory state, e.g., olfaction.

The model presented here is decidedly simple to convey the essence of the problem, and as such, it can be extended in several directions. We would like to note that sensory invariance does not always have to be defined on a single neuron's activity. Any kind of pattern, be that spatial or temporal, can be attempted to be maintained invariant while performing an action. Thus, meaning based on action can also be ascribed to a repeating pattern of activity, not just to a single spike. We believe investigating in this direction will be most fruitful, and in fact we are currently steering our effort into this direction.

6 Conclusion

From the realization that neural decoding methods requiring direct knowledge of the stimulus pose a problem when viewed from within the brain, we derived a novel solution to

the problem of learning the meaning of neural spikes, i.e., through sensory-motor learning based on sensory invariance. We believe that the insight developed in this paper can help build a more autonomous agent with a semantics grounded on its own sensory-motor capacity, for its own sake.

Acknowledgments

References

[1] F. Rieke, D. Warland, R. de Ruter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*, 1st ed.   Cambridge, MA: MIT Press, 1997.

[2] E. D. Adrian, "The impulses produced by sensory nerve endings," *Journal of Physiology (London)*, vol. 61, pp. 49–72, 1926.

[3] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *Journal of Physiology (London)*, vol. 160, pp. 106–154, 1962.

[4] D. Warland and M. Meister, "Multi-neuronal firing patterns among retinal ganglion cells encode spatial information," *Investigative Opthalmology Vision Science Supplement*, vol. 36, p. 932, 1995.

[5] M. W. Oram, P. Földiák, D. I. Perrett, and F. Sengpiel, "The 'ideal homunculus': Decoding neural population signals," *Trends in Neuroscience*, vol. 21, pp. 259–265, 1998.

[6] A. D. Jepson and J. Feldman, "A biased view of perceivers: Commentary on 'observer theory, bayes theory, and psychophysics,'," in *Perception as Bayesian Inference*, D. C. Knill and W. Richards, Eds.   Cambridge University Press, 1996, pp. 229–235.

[7] J. K. O'Regan and A. Noë, "A sensorimotor account of vision and visual consciousness," *Behavioral and Brain Sciences*, vol. 24(5), pp. 883–917, 2001.

[8] S. Hurley, "Perception and action: Alternative views," *Synthese*, vol. 129, pp. 3–40, 2001.

[9] J. J. Gibson, *The Perception of the Visual World*.   Boston: Houghton Mifflin, 1950.

[10] R. A. Brooks, "Intelligence without representation," *Artificial Intelligence*, vol. 47, pp. 139–159, 1991.

[11] P. R. Cohen and C. R. Beal, "Natural semantics for a mobile robot," University of Massachusettes, Department of Computer Science, Tech. Rep. 00-59, 2000.

[12] M. A. Arbib, "Language evolution: The mirror system hypothesis," in *The Handbook of Brain Theory and Neural Networks*, 2nd ed., M. A. Arbib, Ed.   Cambridge, MA: MIT Press, 2003, pp. 606–611.

[13] G. Rizzolatti, L. Fogassi, and V. Gallese, "Neurophysiological mechanisms underlying the understanding and imitation of action," *Nature Reviews Neuroscience*, vol. 2, pp. 661–670, 2001.

[14] T. M. Mitchell, *Machine Learning*.   McGraw-Hill, 1997.

[15] A. J. Bell, "Levels and loops: The future of artificial intelligence and neuroscience," *Philosophical Transactions of the Royal Society of London*, vol. 354, pp. 2013–2020, 1999.

[16] D. Philipona, J. K. O'Regan, and J.-P. Nadal, "Is there something out there? Inferring space from sensorimotor dependencies," *Neural Computation*, vol. 15, pp. 2029–2050, 2003.