Action-Based Autonomous Grounding

Yoonsuck Choe
Department of Computer Science and Engineering
Texas A&M University
College Station, TX 77843-3112, USA

Abstract

When a new-born animal (agent) opens its eyes, what it sees is a patchwork of light and dark patterns, the natural scene. What is perceived by the agent at this moment is based on the pattern of neural spikes in its brain. Life-long learning begins with such a flood of spikes in the brain. All knowledge and skills learned by the agent are mediated by such spikes, thus it is critical to understand what information these spikes convey and how they can be used to generate meaningful behavior. Here, we consider how agents can autonomously understand the meaning of these spikes without direct reference to the stimulus. We find that this problem, the problem of grounding, is unsolvable if the agent is passively perceiving, and that it can be solved only through self-initiated action. Furthermore, we show that a simple criterion, combined with standard reinforcement learning, can help solve this problem. We will present simulation results and discuss the implications of these results on life-long learning.

Introduction

Consider the situation depicted in figure 1. What is shown in these images? There is a black screen, and there are green discs that appear at four discrete locations on the screen. Beyond that, it is hard to even guess what these represent. As unlikely as it seems, this is a typical grounding problem (Harnad 1990), as we can see in the following.
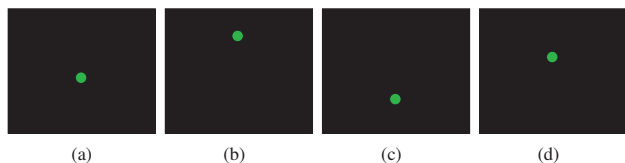

Figure 1: **A Clueless Situation.**

Although the task seems impossible, once the entire context is shown, as in figure 2, now we can see what those green discs (or as it turns out, "neural spikes") represent: They represent visual cortical neurons responding to four different orientations.
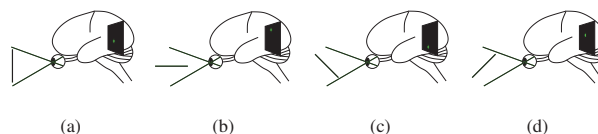
Figure 2: **The Clueless Situation, Explained.**

The situation depicted above seem to indicate that grounding is impossible when only the representations (green discs) are given (as in figure 1). This point can be recast in the context of the brain (figure 3): (a) The second neuron (internal observer) is trying to understand the meaning of spike S it receives (cf. figure 1). (b) An external observer has access to both the spike S and the stimulus I (cf. figure 2).

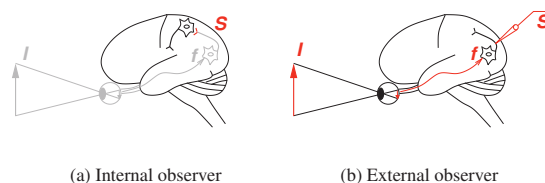(a) Internal observer   (b) External observer

Figure 3: **Internal vs. External Observer.**

A preliminary conclusion from the above would be that there is no way the brain can attach meaning to the spikes (e.g., when a spike fires in response of a diagonal input, it would not know what orientation triggered the spike). However, this conclusion is contrary to our everyday experience since we seem to understand the visual environment without any problem based on just such neural spikes in the brain.

Why were we led to this false conclusion? What is missing from the above? The main ingredient missing from the above (especially figure 3(a)) is the motor system and the sensorimotor feedback. In this extended abstract, we will show how the inclusion of action can help solve this seemingly intractable grounding problem.

Conceptual Solution

First, we consider how, in principle, adding action can help solve the grounding problem shown in the previous section.

We constructed a simple visuomotor agent as shown in figure 4(a) below to investigate the grounding problem. The

agent is situated in a visual environment (left), looking at a small portion of the scene (I). The agent has four orientation sensors (f) which gives the encoded internal state (s). The task is to use action (a) to figure out the meaning of the internal state s by learning a policy $\pi : s \rightarrow a$. This turns out to be a typical reinforcement learning problem. The only thing we need to specify in addition to the above is the reward.

(a) Visuomotor agent
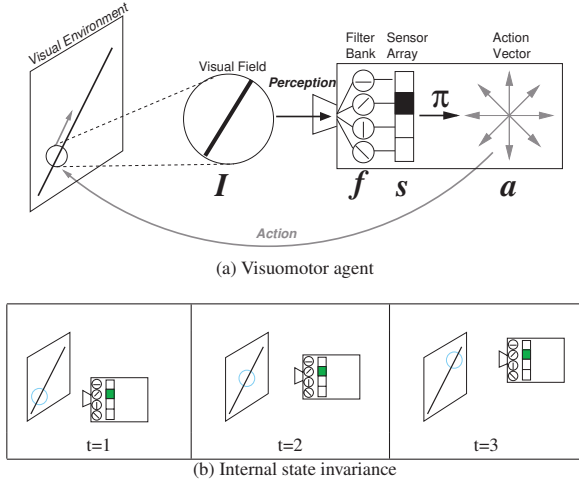

(b) Internal state invariance

Figure 4: **Visuomotor Agent and Internal State Invariance.** Adapted from Choe, Yang, and Eng (2007).

It turns out that a simple criterion can help link spikes to their meaning, in the form of action that carry the same property as the encoded information in the spikes. Figure 4(b) shows how this works. At time $t = 1$, the agent is initially looking at a small part of the diagonal line, and as a result the second unit in its internal state is activated. In the following steps ($t = 2, t = 3$), the agent moves diagonally. However, due to the configuration of the environment and the agent's sensors, the internal state remains invariant (i.e., the second unit is activated). This is strange since movement is expected to lead to a change in the internal state. Here, we can see an interesting correspondence between the invariant state (second unit) and the diagonal movement: The property encoded by the second unit (45° orientation) and the property of the movement (45^0 direction) are the same!

From this simple observation, we can formulate an objective function: given the task above, we can ask the agent to generate action that maximizes invariance in its internal state, from moment to moment. In the following, we will show how this objective can effectively solve the grounding problem.

Experiments and Results

We consider a simple reinforcement learning problem, where the task is to learn the reward table $R(s, a)$, representing $P(a|s)$, to be followed by an optimal policy, where s is the sensory state and a is the action. In an ideal case, the reward table would look like figure 5 below.

We used a simple reinforcement learning algorithm similar to SARSA to estimate the reward table $R(s, a)$ (Choe,

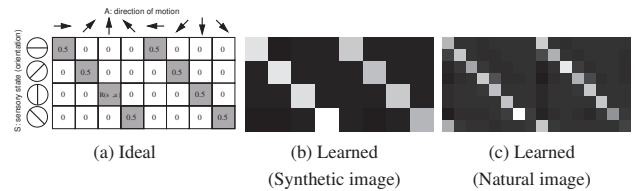| (a) Ideal | (b) Learned (Synthetic image) | (c) Learned (Natural image) |

Figure 5: **Ideal and Learned Reward Tables.** Adapted from Choe, Yang, and Eng (2007).

Yang, and Eng 2007), using synthetic and natural images (e.g., see figure 6(a)). The reward was simply the dot product between the current and the previous sensor response vectors. The learned reward tables are close to the ideal case (figure 5(b)&(c)), which embodies the grounding (synthetic image experiment had 4, and natural image 8 orientation sensors [rows]; columns represent gaze direction). The gaze trajectory based on the learned reward table reveals the underlying structure in the visual scene (figure 6(c)).
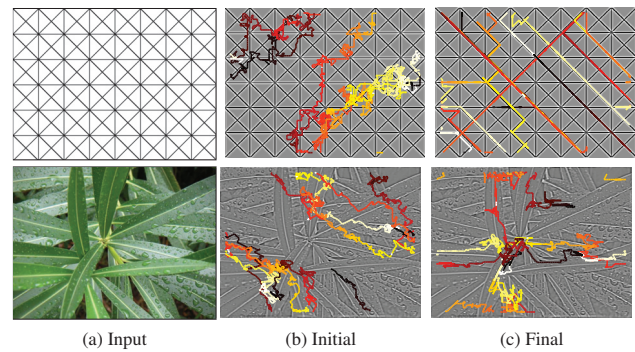

(a) Input   (b) Initial   (c) Final

Figure 6: **Gaze Trajectories.** Color in (b)&(c) indicates flow of time. Adapted from Choe, Yang, and Eng (2007).

Discussion and Conclusion

In this extended abstract, we have shown how action plays a key role in grounding, and identified a simple yet powerful learning criterion: internal state invariance. An important implication of this work is that an agent's "understanding" of the environment critically depends on its motor primitives. This suggests an important requirement for life-long learning: flexible behavior and large motor primitive repertoire. Furthermore, tool use can greatly expand the range of behavior, so tool use can also contribute to understanding and life-long learning.

References

Choe, Y.; Yang, H.-F.; and Eng, D. C.-Y. 2007. Autonomous learning of the semantics of internal sensory states based on motor exploration. *International Journal of Humanoid Robotics* 4:211–243.

Harnad, S. 1990. The symbol grounding problem. *Physica D* 42:335–346.