

Serendipitous Offline Learning in a Neuromorphic Robot

Terrence C. Stewart*, Ashley Kleinhans†, Andrew Mundy‡ and Jörg Conrad§

*Centre for Theoretical Neuroscience, University of Waterloo, Waterloo, ON, Canada

tcstewar@uwaterloo.ca

†Council of Scientific and Industrial Research, Pretoria, South Africa, akleinhans@csir.co.za

Telephone: (800) 555-1212, Fax: (888) 555-1212

‡TODO: Address

Email:

§TODO: Address

Email:

Abstract—We demonstrate a neuromorphic learning paradigm that is well-suited to embodied learning of complex sensorimotor mappings. A mobile robot is first controlled by a basic set of reflexive hand-designed behaviours. All sensor data is provided via a spike-based silicon retina camera (eDVS), and all control is implemented via spiking neurons simulated on neuromorphic hardware (SpiNNaker). Given this initial control system, the robot is capable of simple obstacle avoidance and random exploration. To train the robot to perform more complex tasks, we observe the robot and find instances where the robot accidentally performs the action we wish it to perform. Data recorded from the robot during these times is then used to update the neural control system, increasing the likelihood of the robot performing that task in the future, given a similar sensor state. We demonstrate this learning approach is sufficient for the robot to learn to turn left or right depending on novel sensory stimuli.

Keywords. adaptive systems, mobile robotics, neurocontrollers, neuromorphics, robot control

I. INTRODUCTION

[CHECK: represent related work and compare them with the proposed approach]

A long-standing dream for robotics is to provide the same sort of intelligent, adaptive, and flexible behavior that is seen in living biological systems. Furthermore, by creating systems that emulate biological adaptivity, we can investigate intelligence in a very broad sense, including capabilities that are not yet seen in current machine learning methods [1]. However, it is still a very open question as to how best to make use of this biological inspiration to construct improved methods of robotic control. There are few examples of neural networks being used for control [2] in a robotic domain. Typically, the hardware used is standard general-purpose computing, and the algorithms are machine-learning based. This means that, while they may take some high-level inspiration from biology, the algorithms themselves do not directly map to biological details.

Over the last few years, a bridge has been forming between neuroscience and robotics [3]. Ongoing developments in neuromorphic hardware design now provide novel computing resources that are more suitable to directly implementing

the types of computations found in biological systems, thus facilitating the use of observations and data provided by the neuroscience community. Practically speaking, we can now implement large numbers of neurons on a very small power budget, making neural control a promising energy-efficient direction for mobile robot applications. Ideally, this could lead to flexible adaptive control of systems while staying within a limited power budget.

Our study looks to demonstrate such a flexible control system using neuromorphic hardware and neural-based adaptive control. The approach combines the SpiNNaker computing platform [4], [5], the Neural Engineering Framework (NEF) [6], and a series of robots developed at the Technische Universität München.

For energy efficient implementation given the hardware, the neuron model employed here is the standard *leaky integrate and fire* (LIF) model, where the weighted (ω) sum of input current I causes the voltage V to build up until some threshold is reached, at which time the neuron emits a discrete spike:

$$\tau_m \frac{dV}{dt} = -V + \sum \omega I \quad (1)$$

When this spike occurs, this causes post-synaptic current to flow into the neurons it is connected to, with each connection having its own weight ω and with the post-synaptic current $h(t)$ exponentially decaying over time:

$$h(t) = e^{-t/\tau_s} \quad (2)$$

The goal of this work is to explore algorithms that can be usefully implemented given hardware that efficiently implements neural networks of this form. In particular, we note that living creatures have both genetic, low-level hard-wired reflexes and they are also capable of developing novel associations between stimuli and responses that are entirely context-dependent. Behavioural studies indicate that they can learn to perform certain actions at certain times, through experience, overriding and building upon these low-level reflexes [7].

However, for applied robotics applications, we do not want an approach where learning is entirely autonomous and

Fig. 1. The PushBot and the eDVS silicon retina.

undirected (as in, for example, the neural learning seen in Distributed Adaptive Control [8]). Instead, our approach is to inform neural learning by providing explicit indications as to the instances where correct behaviour was achieved. The approach described here is somewhat akin to reinforcement learning, but relies only on positive examples, and can be explicitly shaped as desired. This guides the learning and provides explicit control over the eventual behaviour.

II. INFRASTRUCTURE

[CHECK R3: "proposed approach is not well organised and written, therefore cannot see valuable novelty in manuscript."]

A. eDVS

The sensor system used here is the eDVS embedded dynamic vision sensor [9], a silicon retina developed by iniLabs in collaboration with the Institute of Neuroinformatics at the University of Zurich and the ETH Zurich. This neuromorphic sensor is a 128x128-pixel camera. Instead of reporting frame-based data, it emits individual events when the relative brightness for any individual pixel increases or decreases. The eDVS provides high temporal resolution (1 μ s), low latency (15 μ s) and high dynamic range (120dB). The eDVS used here is an embedded version with an onboard microcontroller (NXP LPC4337), inertial measurement unit, multiple PWM control signals, and general-purpose I/O lines. The silicon retina is well-suited for integration with other neuromorphic hardware since it produces output in the form of spikes, which is the natural communication framework for spike-based neural processors. Furthermore, certain image processing algorithms can be implemented efficiently. For example, previous work [10] has implemented high-speed tracking of flashing LEDs, and we use that algorithm here as sensory pre-processing.

B. PushBot

At the Technische Universität München, the Neuroscientific System Theory group has developed a small tread-based robot built around the eDVS sensor, as shown in Fig. 1. The robot provides two motors, a frequency-controllable laser pointer, two LEDs, and housing for power (4 AA batteries). The setup comes with a proprietary WLAN module, enabling streaming of sensor data and motor command to and from the robot over standard WiFi.

C. SpiNNaker

The SpiNNaker multicore processor is developed by the University of Manchester and consists of 18 200MHz ARM968 cores on a single die [4], [5]. The processors and inter-chip communication infrastructure are optimized for massively parallel operations, with a target of a one-million-core machine. Its low power consumption means that a 48-chip SpiNNaker board with a total of 864 processors uses under 40W of power. This system can be programmed directly in C or using the standard neural modeling API PyNN.

However, for this work we made use of the Nengo open-source neural compiler [11] (described below), and the custom Nengo backend that compiles high-level neural models into optimized SpiNNaker code [12].

D. Nengo and the Neural Engineering Framework

The reason we chose Nengo rather than a more standard neural network framework is that Nengo [11] directly instantiates the Neural Engineering Framework (NEF).

The Neural Engineering Framework (NEF) is a general-purpose neural compiler that allows the user to define a high-level algorithm that is then compiled down to a neural approximation of that algorithm [6]. This approach was originally meant for constructing complex biologically realistic neural models. Demonstration of the use of Nengo for perception, cognition, and action encoding can be seen in Spaun, the first large-scale (2.5 million neuron) functional brain model capable of performing multiple tasks [13].

When developing neural models using the NEF, every group of neurons is thought of as representing a vector. That is, different patterns of activity within that group of neurons correspond to different vector values being represented. This vector is (normally) of much smaller dimensionality than the number of neurons in the group, so the neural firing forms a redundant code for that vector. Each neuron in the group responds differently to the same input, due to neural heterogeneity. Therefore, the redundant code can be thought of as a random projection from a low-dimensional space (the vector being represented) to a high-dimensional space (the neural activity).

Connections between neural groups implement functions on those represented vectors. Importantly, due to the redundant code, these functions do not have to be linear functions; rather, any function can be specified, and the system will find the optimal connections between neurons to approximate that function. The Neural Engineering Framework treats this as a least-squares minimization problem and finds the feed-forward linear synaptic connection weights between the individual neurons that will most closely approximate the desired nonlinear function. Furthermore, the NEF also indicates how recurrent connections can be found that will approximate any desired differential equation. For both feed-forward and recurrent connections, the accuracy of the approximation will be dependent on the number of neurons and the function being approximated.

As an example, consider a group of neurons storing three values: the x, y position of an object and a certainty measure c that the object is at that location. (This example comes from the output of the tracking algorithm for flashing LEDs [10]). In the NEF, we might use 100 neurons to form a distributed representation of these three values (using more neurons would improve the accuracy of the representation). Each neuron would be given a random combination of inputs from the three dimensions, resulting in a different pattern of firing for different vectors (x, y, c) .

Given this activity, we can then define connections to other groups of neurons that compute functions of these values. For example, if we want a group of neurons R to store the distance from the center of the visual field to the location of the LED, we could tell Nengo to optimize the connections so as to approximate:

$$R \leftarrow \text{sqrt}(\text{led.x}^2 + \text{led.y}^2)$$

This would find synaptic connection weights between the group of neurons representing the LED data and the group of neurons representing the radius R such that R is driven to fire with whatever pattern represents the correct radius, given the current activity of the LED population.

All of the neural systems described below are defined in this manner.

III. METHOD

The methodology used here is to start by building a neural system that implements a base set of automatic “reflex-like” behaviors for the robot. This is meant to correspond to the hardwired automatic responses found in living creatures. In addition to this, we will then add a learned system where the robot must map its sensory state to the particular desired actions for this condition. This corresponds to learned associative responses in animals.

[TODO R2: provide more insight into methodology, to reflect scientific paper rather than technical report]

A. Initial Reflexive Control

The first stage of this study requires a definition of the base set of simple behaviors and their triggering conditions, for a small mobile robot. These behaviors should be as simple as possible, since they must be hand-designed, but should provide a fairly rich repertoire of resulting actions. These can be thought of as, the basic, instinctive reflexes seen in living creatures. In general, the first step in using this methodology is to define a collection of these reflexes.

The first reflexive behavior is simply to go forward if there is no obstruction. To define this, we specify a function that uses sensor data to compute the current strength S of an action (i.e. 0 to 1). In this case, the visual position of the dot caused by the laser pointer mounted on the robot can be used as a simple range detector: if it is lower than some threshold, the robot is near a wall and should not go forward. If it is higher than that threshold, it is safe to go forward:

$$S[0] \leftarrow \text{if } \text{laser.y} > -0.6 \text{ else } 0$$

To complete this basic behavior, we define the motor output M as a function of S . In this case, we want to drive both wheel motors forward when the action has a high strength.

$$M \leftarrow [1, 1] * S[0]$$

The next basic reflexive action is to back up when we are too close to an obstacle. This is again detected by using the visual location of the dot shown by the laser pointer. If this is very low in the visual field, or if it is not visible at all (i.e. the

dot is below the field of view of the camera) then the robot should back up. This gives the following rule:

$$S[1] \leftarrow 1 \text{ if } \text{laser.y} < -0.8 \text{ or } \text{laser.c} < 0 \text{ else } 0 \\ M \leftarrow [-1, -1] * S[1]$$

The final basic actions are: turn left or right when close to an obstacle.

$$S[2] \leftarrow 1 \text{ if } \text{laser.y} < -0.8 \text{ else } 0 \\ M \leftarrow [-1, 1] * S[2]$$

$$S[3] \leftarrow 1 \text{ if } \text{laser.y} < -0.8 \text{ else } 0 \\ M \leftarrow [1, -1] * S[3]$$

However, these last two actions should not both be performed at the same time, since this would cause the robot to stay in place, rather than turning one direction or the other. To specify this, we can define connections that relate the strengths of different actions to each other.

$$S[2] \leftarrow -S[3] \\ S[3] \leftarrow -S[2]$$

This means that whenever $S[2]$ is positive, it will force $S[3]$ towards 0, and similarly $S[3]$ will drive $S[2]$ towards 0. This implements a classic neural competition, meaning that only one of the two actions will have a large value at a particular time.

Now that these reflex actions have been defined in these simple terms, we use Nengo to build neural networks that approximate each of these rules. Groups of neurons are defined to store the sensory state, the motor output, and the strengths of each action. The connections between the groups of neurons are set to approximate each of the above functions, resulting in the neural system shown in Fig. 2.

To understand the resulting basic reflexive behavior, two important aspects of NEF must be considered. First, due to the neural representation scheme, when there are multiple inputs to a neural population, the value represented will be the sum of the input values. In other words, the motor output will be the sum of all the connections: $S[0]$, $S[1]$, $S[2]$, and $S[3]$ to M . Second, the NEF approximations of the desired functions will be *smooth*. That is, instead of implementing the exact functions specified above, the actual behavior will be much more gradual. For example, the input function to the $S[0]$ population is supposed to compute 4 if $\text{laser.y} > -0.6$ else 0. Instead of implementing this exact function, the neurons will instead implement a smoothed approximation of this step function, as shown in Fig. 3.

The result is a small spiking neural network control system that performs very simple obstacle avoidance. Due to the smoothing inherent in the NEF neural approximation of the above functions, the robot gradually transitions between behaviors. This means that it will automatically slow when approaching a wall, even though we have not explicitly built reflexes to do so.

When it reaches a wall, it will then turn in a random direction. This randomness will be due entirely to sensory

Fig. 2. [TODO R1: Terry fig’s 5,6 and 7 should all show 10 trials (ie same amount of trails)] Basic reactive control. Each box is implemented as a group of leaky integrate and fire neurons, where N indicates the number of neurons, whose tuning curves span the dimensionality D as indicated. Each solid arrow indicates an alltoall connection with connection weights computed as per the NEF to approximate the functions described above. Dotted arrows indicate inputs and outputs connecting the robot to the neuromorphic hardware.

Fig. 3. Neural approximation with the NEF. When connections between neural groups are optimized to approximate a function, the result is a smooth version of that function. As the number of neurons is increased, this neural approximation will approach the ideal function.

Fig. 4. The T-Maze environment. The robot starts at the bottom.

Fig. 5. Behavior after learning to turn left. By adding connections optimized to approximate situations where the robot behaved appropriately, we implicitly program the robot to map its sensory states to its actions as desired.

noise, as there is no stochasticity in the neural model itself. Once it has chosen a particular direction to turn, i.e. once either $S[2]$ or $S[3]$ has a high value, it will continue to turn in that direction until there is no longer an obstacle in front of it. All actions transpire without human interference. The only hardcoded input are the initial thresholds that provide information as to whether an action is, in fact, correct given the current sensory state. These actions demonstrate obstacle avoidance, and simple decision making, in much the same way as animal behavior studies [7].

The core idea is that this set of behaviors is a bare minimum, sufficient to allow the robot to move throughout its environment without human intervention. As with animal behaviour, this allows the robot to have a basic set of automatic decision-making rules that apply in novel situations where there is not yet any learned response. Using these reflexive rules, the robot makes “accidental” decisions (turning left in some situations and turning right in others, for example), which we will then use to further train the robot’s behavior, as discussed in the next section.

B. Initial behavior

To examine the model’s behavior, we used a standard T-maze environment (Fig. 4). When placed at the bottom of the T-maze, the robot navigates forward, avoids the walls, reaches the choice point, and turns left or right. Typical trajectories are shown in Fig. 1, which indicates the motor outputs over time. Since the robot uses tank-style treads, the two motor output values are the left and right motor speeds, respectively. For clarity, here we plot the overall speed as the sum of $M[\text{left}]$ and $M[\text{right}]$, while the rotation rate is the difference between the two. Motor output values of typical individual runs are plotted along with an overall mean value (with 95% bootstrap confidence interval). All values are Gaussian filtered with $\sigma = 0.15s$ for visual clarity.

C. Serendipitous Offline Learning

While the above approach is sufficient for implementing simple behaviors, it requires the developer to define explicit

Fig. 6. behavior of reactive control model over multiple runs. The speed (top graph) is high at the beginning, then slows as it turns either left or right (bottom graph). While on any individual run the robot tends to turn consistently either left or right, the overall average is zero turning (black area in bottom graph).

rules mapping sensory states to actions. In this section, we build upon these rules in a less explicit manner, allowing us to create more complex behavior without hard-coding any additional rules. In complex situations, it may not be feasible for action rules, such as the ones defined above, to be hard coded. Instead, we can also define *implicit* action rules. The basic idea is to allow the robot to explore its environment, but whenever it happens to accidentally do whatever action we want it to do, we record the sensory and motor values and *use those values to define a new rule*.

For example, consider the simple case where we want the robot to always turn left at the intersection, rather than randomly turning either left or right. Recording the sensory data and the strength of each action S while the robot is performing its initial basic behavior, we find instances where the robot performs the desired action (turning left). We call these the *positive examples*. In this case, we consider the four individual runs shown in Fig. 1 [TODO: check all figures are added and that they refer to the correct figures] where the robot’s rotation tended to remain positive. The data from runs where it turned right are removed. The data from these positive examples can be thought of as long sequences of state-action pairs, indicating what action to perform in what state. Given this, we add new connections from sensors to the action strength populations. Instead of explicitly indicating what functions to approximate on these connections, we instead use the data gathered from the robot itself to be the target function that the neurons must approximate.

At first glance, it seems as if these new connections would just end up being exactly the same as the original reflexive control connections. However, the key difference here is that these new connections will also take into account *other sensory states*, not considered in the original hand-designed reflexive rules. In other words, these new connections will cause the robot to be more likely to perform the actions we desire whenever it is in a sensory state similar to those seen in the positive examples. Importantly, this will happen without any explicit indication of exactly what sensory states should trigger what actions. This allows the system to discover more complex rules than could be reasonably manually hand-designed.

IV. RESULTS

The result of this training is shown in Fig. 5. Unlike Fig. 6 this shows the robot consistently turning to the left about two seconds into the task (the average time for the robot to reach the end of the corridor).

A. Learning Sensory Conditions

The initial example of learning to turn left is not particularly complex. However, we can use exactly the same process to

Fig. 7. behavior after learning to turn right if there is a mirror, and otherwise turn left. The robot successfully identifies the correct situation and turns appropriately. Robot speed is not shown, but is similar to that depicted at the top of Fig. 6

produce more complex behavior. To demonstrate this, we now sometimes place a *mirror* at the intersection. Initially, the robot will ignore the mirror and just use its standard random reflexive navigation. We now choose as our positive examples situations where the robot turned right when there was a mirror, and situations where it turned left when there was no mirror.

Adding a learned connection that attempts to approximate those positive examples means that the neural connections now need to change their outputs depending on whether the mirror is present or not. Fig. 7 shows that the robot can successfully learn to recognize and respond to mirrors, given only this simple learning rule on top of its basic reflexive rules.

V. DISCUSSION

The algorithm described in this paper is meant to be a general-purpose approach, making use of massively parallel low-power neuromorphic computing hardware on a mobile robot. We have shown that we can start by defining extremely simple reactive rules, using Nengo and the NEF to implement them on neuromorphic hardware. Certainly, these sorts of simple reactive rules could also be implemented using more traditional computing hardware. However, the point of this paper is to demonstrate automatic training of a simple set of neural connections that interacts with basic reactive systems to produce much more complex control.

In the case of the system learning to change its behavior based on the presence of a reflective mirror surface, the system was able to map the complex sensory stimuli to a particular motor action. In particular, it was able to make use of whatever sensory discrimination is available via the mirror (or lack of a mirror). It is important to note that this sensory stimulus did not initially impact the robot's behavior in any way. It was only through the process of learning, i.e. using hindsight examples of desired behavior and building a new set of connections that approximate required behavior that triggered changes in movement. This was done by building neural connections that replicate behaviour that the robot previously did accidentally.

This means that the scope of possible sensorimotor mappings is limited only by the range of the sensory inputs and the size of the neural population representing the sensors. If there are sufficiently many neurons representing the sensory space, then NEF guarantees that connection weights can be found with a desired level of accuracy, if there is some function that could map the sensory state to the motor action. However, for complex functions, the number of neurons required may be excessive. Thus, further investigation is required to characterize a variety of different learned tasks with various sensor systems. This would indicate what types of actions are feasible to learn in this manner.

This general approach of using sensory experience to learn neural connections that attempt to cause the same actions to occur in response to similar sensory stimuli is novel, but we have performed previous work [14] with a somewhat similar learning method. In that case, we did not have an initial reflexive control system, and we did not find particular examples of desired behavior. Instead, we put the robot temporarily under manual control and used the behavior generated by the person performing the manual control to train the connections between sensory neurons and motor neurons. In other words, rather than allowing the robot to randomly explore its environment and choose actions based on a manually created reflexive control system, we had a user manually trigger actions using a remote control. As in this work presented here, we then trained neural connections that would approximate that manual control.

In that earlier work, learning required direct training examples, rather than simply labelling particular actions as positive example of desired behavior after they occur. Indeed, this means that the work presented here could be thought of as reinforcement learning, while the previous work would be purely supervised learning. Of course, a hybrid approach could be pursued.

Finally, it should be noted that we are only training based on positive examples (i.e. situations where the robot behaved as desired). In future work, we plan to augment this approach with explicit punishment for situations where the robot performs a non-desired behavior. Interestingly, adding this capability is not straight-forward in this case. Indeed, there is significant biological evidence that positive (reward), and negative (punishment) systems are separate in the brain [TODO: reference]. That is, they are not simply the opposite of each other – rather they are significantly different processes that interact. This interaction is still to be explored. One step towards this is a model of fear conditioning [15], which would fit well with this framework.

ACKNOWLEDGEMENT

The Telluride Neuromorphic Cognition Workshop 2011.

REFERENCES

- [1] D. McFarland and T. Bösser, *Intelligent behavior in animals and robots*. Mit Press, 1993.
- [2] D. Janglová, "Neural networks in mobile robot motion," *Cutting Edge Robotics*, vol. 1, no. 1, p. 243, 2005.
- [3] B. Webb, "What does robotics offer animal behaviour?" *Animal behaviour*, vol. 60, no. 5, pp. 545–558, 2000.
- [4] S. Furber and S. Temple, "Neural systems engineering," *Journal of the Royal Society interface*, vol. 4, no. 13, pp. 193–206, 2007.
- [5] S. B. Furber, F. Galluppi, S. Temple, L. Plana *et al.*, "The spinnaker project," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 652–665, 2014.
- [6] C. Eliasmith and C. H. Anderson, *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. MIT press, 2004.
- [7] Y. B. Kim, N. Huh, H. Lee, E. H. Baeg, D. Lee, and M. W. Jung, "Encoding of action history in the rat ventral striatum," *Journal of neurophysiology*, vol. 98, no. 6, pp. 3548–3556, 2007.
- [8] P. F. Verschure, "Distributed adaptive control: a theory of the mind, brain, body nexus," *Biologically Inspired Cognitive Architectures*, vol. 1, pp. 55–72, 2012.

- [9] J. Conradt, R. Berner, M. Cook, and T. Delbruck, "An embedded aerodynamic vision sensor for low-latency pole balancing," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 780–785.
- [10] G. R. Müller and J. Conradt, "A miniature low-power sensor system for real time 2d visual tracking of led markers," in *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2429–2434.
- [11] T. Bekolay, J. Bergstra, E. Hunsberger, T. DeWolf, T. C. Stewart, D. Rasmussen, X. Choo, A. R. Voelker, and C. Eliasmith, "Nengo: a python tool for building large-scale functional brain models," *Frontiers in Neuroinformatics*, vol. 7, Jan. 2014. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3880998/>
- [12] A. Mundy, J. Knight, T. C. Stewart, and S. Furber, "An efficient SpiNNaker implementation of the Neural Engineering Framework," 2015.
- [13] C. Eliasmith, T. C. Stewart, X. Choo, T. Bekolay, T. DeWolf, Y. Tang, C. Tang, and D. Rasmussen, "A large-scale model of the functioning brain," *Science (New York, N.Y.)*, vol. 338, no. 6111, pp. 1202–1205, Nov. 2012.
- [14] J. Conradt, F. Galluppi, and T. C. Stewart, "Trainable sensorimotor mapping in a neuromorphic robot," *Robotics and Autonomous Systems*, 2014.
- [15] C. Kolbeck, T. Bekolay, and C. Eliasmith, "A biologically plausible spiking neuron model of fear conditioning," *Robotics and Autonomous Systems*, pp. 53–58, 2013.