Stability of Learning in the ARBIB Autonomous Robot

R. L. B. French **R. I. Damper**
Image, Speech and Intelligent Systems (ISIS) Research Group,
Department of Electronics and Computer Science,
University of Southampton,
Southampton SO17 1BJ, UK.

Abstract

We have previously described the ARBIB autonomous robot which consists of a mobile platform running a neural network simulator. Unlike most other behaving robots, the neural system is biologically-inspired and operates at the level of individual spikes. Rather than using currently-popular reinforcement learning techniques, ARBIB learns from exposure to its environment via low-level mechanisms of habituation, sensitisation and classical conditioning. In previous work, its short-term memory (formed through recovery of synaptic weights) made its learning almost totally plastic; it had no long- or medium-term memory. This paper explores means of adding stability to its learning.

Long-term memory is provided by a simple form of synaptogenesis, which forms new connections within the nervous system. Medium-term memory is provided by a recurrent neural circuit coupled to a simple model of the medial pallium, which in turn is fed from a sonar range-finding cell. This allows the nervous system to respond to successive stimuli that lie outside the duration of its short-term memory.

The effects of these two enhancements are assessed for their impact on ARBIB's behaviour. Long-term memory is tested by collision avoidance behaviour, demonstrated by presenting a comparison of firing activity in bump sensory cells with and without synaptogenesis. Medium-term memory is tested by allowing the sonar-driven medial pallium to habituate to a distant target, and then introducing a transitory target at close range. In this way, a useful measure of learning stability through medium- and long-term memories is achieved.

1. Introduction

Previously, we have described ARBIB: an **A**utonomous **R**obot **B**ased on **I**nspirations from **B**iology (Damper, French, and Scutt, forthcoming). In engineering terms, it consists of a mobile platform (a Nomad Scout in this work) running a neural network simulator called Hi-NOON. Unlike other behaving robots, the neural system is biologically-inspired and operates at the level of individual spikes. Rather than using currently-popular reinforcement learning techniques, ARBIB learns from exposure to its environment via low-level mechanisms of habituation, sensitisation and classical conditioning. We see ARBIB as a vehicle for studying "the animat path to AI" (Wilson 1991).

One of the central 'tenets of good old fashioned AI' (Haugeland 1985) was implementation independence. That is, intelligence is achieved by running some 'program'; the hardware on which it runs is unimportant. By contrast, this work is motivated by new ideas of 'embodied AI' which hold that possession of a body situated in its environment is a vital part of being a behaving organism (Churchland 1986; Clark 1987; Brooks 1991). Further, any such organism has a nervous system which obviously underlies its intelligent behaviour. Philosophically, then, we are attracted to the *neuron doctrine* of Barlow (1972), according to which behaviour is inextricably linked to neurophysiology. This is the reason for the biological fidelity (spiking neurons and neurobiological learning mechanisms) with which ARBIB's nervous system is modelled/implemented.

As previously implemented, however, ARBIB suffered a major shortcoming in that its learning was far too plastic (because of recovery of synaptic weights). This problem has been well recognised in the *stability-plasticity* dilemma of Grey Walter (1951, p. 63) and Carpenter and Grossberg (1988). That is: How can a nervous system retain its stability of learning while still being plastic enough to adapt in a changing environment? Consistent with the philosophy outlined above, we seek to add stability (in the form of medium- and long-term memory) by biologically plausible means.

The remainder of this paper is structured as follows. In Section 2, we give an overview of ARBIB's neural simulator, called Hi-NOON. Section 3 gives the equations governing nervous system behaviour. Section 4 introduces the biological phenomenon on which we have based ARBIB's long-term memory: synaptogenesis. It is used in simplified form in this work. We also present results of testing this long-term memory with obstacle avoidance experiments. Section 5 describes implementation of a form of medium-term memory. Section 6 briefly compares ARBIB with other animat studies. Finally, Section 7 concludes with some suggestions for future work.

2. Overview of Hi-NOON

In this section, we describe the neural simulator Hi-NOON, which stands for **Hi**erarchical **N**etwork of **O**bject-**O**riented **N**eurons. As the name suggests, in Hi-NOON, synapses, neurons and networks are in principle represented as objects within an object-oriented hierarchy (Scutt and Damper 1991, 1997) at various levels of abstraction. The lowest such level uses the membrane potential (strictly, transmembrane potential difference) as the observable parameter in the network model. Consistent with our underlying philosophy, this is a much lower-level approach than the use of activation values roughly corresponding to the spike or action potential rate of individual neurons or collections of neurons as in parallel distributed processing (PDP) models. By contrast, Hi-NOON retains details of individual spike generation which is lost in the traditional connectionist approach. Also, Hi-NOON facilitates simulation of a non-homogeneous population of neurons.

2.1 *Neuron Parameters*

Basic neurophysiology suggests the attributes a model spiking neuron should have. The fixed parameters `BaseMP`, `Threshold` and `TimeConst` correspond to the resting potential, threshold and time constant of the neuron, respectively. Dynamic parameters `MP`, `SynPot` and `fired` (a 1/0 predicate) model the actual membrane potential as it varies in time, accumulate the weighted sum of synaptic inputs which influence the updating of `MP` at the next time step, and indicate if the object is in the process of firing, respectively. This parameter system allows us easily to describe differences between neurons and to keep track of the changing states of neurons over time. It approximately satisfies the “minimum requirements” for effective neural modelling detailed by Selverston (1993).

2.2 *Hi-*NOON *Objects*

The neural network is represented as a list of objects, where each such object corresponds to a single neuron and holds all the information about its state (see below) and about subsidiary objects. The neuron object comprises information on:

- a set of parameters which defines the neuron;
- a set of data structures which defines the ‘axon terminals’ for the neuron, each of which is itself an object and has its own parameters;
- a set of methods – pointers to functions – which access and alter parameter values and so determine exactly how the neuron functions.

The top-level list corresponds to the network object. This possesses two methods – `h_access` and `add` – for accessing network objects and adding further objects onto the list, respectively. Simulation run length is handled by a global object. This stores the simulation and concurrent socket interface ‘housekeeping’ data, including a counter whose original value specifies the length of simulation. It decrements after each evaluation of the network object, and the simulation halts when the counter reaches zero.

As synapses are also objects, they too have fixed and dynamic parameters similar to those of neurons. Thus, `BaseWeight` is the default weight of the synapse and is a constant; `Weight` holds the present synaptic strength and is variable during simulation; `Recovery` is a constant (within each synapse) which determines how quickly `Weight` returns to `BaseWeight`. To prevent synaptic weights growing without limit, `Weight` is bounded during simulation. This models the finite stores of neurotransmitter in the synaptic terminals of biological neurons.

2.3 *Neuron Types*

Hi-NOON allows a non-homogeneous population of neurons to be simulated – reflecting the fact that neurons have specialised functions in real neurobiological systems – at the most appropriate level of abstraction. Modeling individual neurons at the level of membrane potential allows subthreshold and spiking behaviours to be simulated at low computational cost. The fixed parameters cater for differences between neurons which, in this work, are of the following types:

basic: tells its synapses to fire when its membrane potential crosses threshold from below.

noisy: similar to **basic**, but has an additional internal noise component determining the weighted synaptic input, and hence influencing the membrane potential at the next time step.

ramp: similar to **noisy**, but has ability to ramp up spike generation rate. It is used as a test signal source in network development.

burst: similar to **noisy** but produces a short burst of spikes when its membrane potential crosses threshold.

sensor: similar to **basic**, but acts as a sensory neuron in a situated system, such as a mobile robot.

motor: similar to **basic**, but acts as a motor neuron in a situated system.

2.4 *State System*

Each neuron is treated as being in one of a number of six states depending on the present membrane potential, cell threshold and whether or not the cell has just fired, etc. For example, if the membrane potential of the basic cell is above threshold and the cell has not just fired, then the neuron will start to generate a spike and will initiate synaptic transmission.

Figure 1 (taken from a Hi-NOON simulation) shows the states passed through by a neuron during firing. In the case

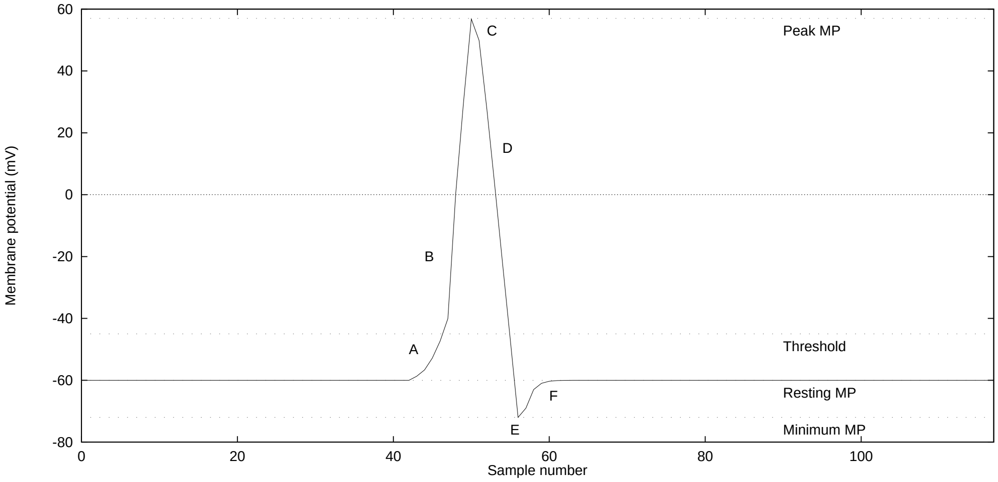


Figure 1: Time evolution of typical action potential (spike) of a **basic** neuron in a Hi-NOON simulation. See text for specification of the states (A..F) passed through by a neuron during firing. Here, the sample period is approximately 4 ms (this varies with the machine on which the simulation runs.)

illustrated, the minimum, resting and peak potentials of the neuron are set at −69, −60 and +45 mV respectively, and the threshold value was −45 mV. (Note that actual values will under/overshoot these settings before state can change at the next iteration of simulation.) The states are:

A:	`MP` above resting potential and below threshold
B:	above threshold and below peak
C:	at peak
D:	post-firing
E:	at minimum
F:	hyperpolarised

Equations governing the membrane potential in each of these states and behaviour of synaptic weights are given in Section 3.

The use of a state system for controlling the membrane potential facilitates the addition of new features to the program; it is only necessary to identify which of the states may trigger this feature and to add a procedure call at that particular state. This, coupled with the inheritance mechanism of object oriented programming (Coad and Yourdon 1991), allows models to be developed and altered with a minimum of changes to the neural model source code.

2.5 Axonal and Synaptic Transmission

We model both sub- and supra-threshold behaviour of neurons, but sub-threshold potentials are not propagated (from axon hillock to terminal fibres) in real neurons, only action potentials are. We do not attempt to model (regenerative) spike transmission along the axon. This, however, is not a serious concern because the model's behaviour depends entirely on how pre-synaptic activity is transformed into post-synaptic activity. It is only in supra-threshold states B, C and D (Figure 1) that synaptic communication can take place. Hence, it is irrelevant that we are, in some sense, modelling sub-threshold behaviour incorrectly. An alternative view is that we are not modelling axonal transmission, i.e. we have 'point' neurons as is common in neural modelling (MacGregor 1987, pp. 21–24).

*2.6 Learning in Hi-*NOON

There is no specific support for learning in Hi-NOON. Thus, if PDP-type learning (e.g. back-propagation) is to be used, this must be implemented external to the simulator. In light of Hi-NOON's ability to model at the level of transmembrane potential, however, there is implicit support for biologically-based forms of learning operating at the neuron level. In an intriguing paper, Hawkins and Kandel (1984) argue that the cellular mechanism underlying classical conditioning in the well-studied sea slug *Aplysia* is an extension of the mechanism underlying sensitisation. They go on to state [p. 375]: "This finding suggests that the mechanisms of yet higher forms of learning may similarly be based on the mechanisms of … simple forms of learning" referring to this as "an elementary cellular alphabet of learning" [p. 376]. Thus, we build classical conditioning on the basic neural model of habituation/sensitisation. Generally, these simple forms of learning are implemented using synapse-on-synapse connections in Hi-NOON as described below.

3. Neurons and Synapses

In this section, we present more detailed descriptions of neurons and synapses within Hi-NOON. Since it is intended for (among other things) applications in situated robotics studies, there is provision for sensory and motor neurons which connect to the environment, as well as for more 'basic' (information processing) neurons.

3.1 Neurons

The ‘basic’ neuron type has the state system functionality which is subsequently embedded in all derivatives, such as the sensory and motor cells.

3.1.1 Basic neurons

Updating equations for the membrane potential (MP – in millivolts) for this neuron type are:

$$\begin{array}{ll}
\text{state A:} & MP(t+1) = MP(t) - \tau + S(t) \\
\text{state B:} & MP(t+1) = MP(t) - \alpha + S(t) \\
\text{state C:} & MP(t+1) = h + S(t) \\
\text{state D:} & MP(t+1) = MP(t) - \mu + S(t) \\
\text{state E:} & MP(t+1) = l + S(t) \\
\text{state F:} & MP(t+1) = MP(t) + \frac{BaseMP - MP(t)}{\eta} + S(t)
\end{array}$$

where:

$S(t)$ is the synaptic potential (`SynPot`) at time t, equal to $\sum_i w_i \kappa \left(MP_i(t) - BaseMP_i\right)$;

i is a counter which counts over active pre-synaptic cells;

w_i is the synaptic weight from a pre-synaptic neuron;

τ is the neuron time constant;

$\eta = 1.5$ is the post-undershoot increment rate;

$\mu = 25$ is the post-action potential peak-MP decrement;

$\kappa = 1/450$ is a heuristically-set learning constant;

$\alpha = 20$ is the post-threshold attack increment;

$h = 45$ is the post-threshold maximum MP;

$l = -69$ is the pre-undershoot minimum MP;

Certain of these parameters (e.g. τ, η) are time-dependent and have been set empirically to suit a range of processor speeds and implementations. However, they may be inappropriate in some circumstances (as when implementing a real-time robotic system using a fast processor).

3.1.2 Sensory Neurons

The sensory neurons in this work react to object proximity. Distance sensors feed sensory cells with a sampled value (approximately) proportional to object range, $R(t)$, whence:

$$MP(t+1) = MP(t) + \lfloor sR(t) \rfloor$$

where $\lfloor\ \rfloor$ represents the floor function and s is an empirically set scaling factor whose value depends on the particular physical sensor employed.

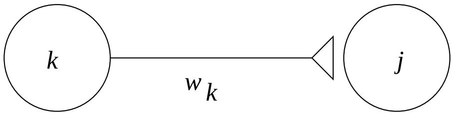


Figure 2: Synapse from parent neuron k to target neuron j, with weight w_k.

3.1.3 Motor Neurons

These close the loop between the nervous system and the environment. Motor drive activity is given as:

$$A(t) = \left\lfloor \frac{\pm\gamma MP(t)}{h} \right\rfloor$$

where $A(\,)$ is used at a lower level of abstraction (closer to the hardware) to determine the speed of motor output, and γ is a scaling constant set to suit the particular robot hardware in use. Its sign is determined by the requirement for forward or reverse drive.

3.2 Synapses

The basic synapse (which is noise free) is shown in Figure 2. It has functionality which is subsequently embedded in all derivatives such as the habituating, sensitising and conditioning types. These allow us to implement a simple, biologically-based form of learning.

$$w_k(t+1) = \begin{cases} w_k(t) - \beta & \text{if } w_k(t) > w_{\text{base}} \\ w_k(t) + \beta & \text{if } w_k(t) \le w_{\text{base}} \\ w_{\max} & \text{if } w_k(t) > w_{\max} \\ w_{\min} & \text{if } w_k(t) < w_{\min} \\ w_{\min} & \text{otherwise} \end{cases}$$

where β is the MP recovery parameter and w_{base} is the base weight (typically 0). These are individually set (together with $w_{\min}$ and $w_{\max}$, typically ± 16) for each neuron.

3.2.1 Noise-free synapse

A noise-free synapse is guaranteed to fire whenever the parent neuron fires, according to:

$$\texttt{fired}(t) = \begin{cases} \text{TRUE} & \text{if state B, C, D} \\ \text{FALSE} & \text{otherwise} \end{cases}$$

3.2.2 Noisy synapse

A noisy synapse will fire, with a known probability, whenever the parent neuron fires:

$$\texttt{fired}(t) = \begin{cases} \text{TRUE} & \text{if } \texttt{cond} \\ \text{FALSE} & \text{otherwise} \end{cases}$$

where p denotes a parent (pre-synaptic) neuron, θ is its threshold, and:

$$\texttt{cond} = (\text{state B, C, D}) \wedge \left(\frac{MP_p - \theta_p}{h - \theta_p} \ge \text{rand}(1) \right)$$

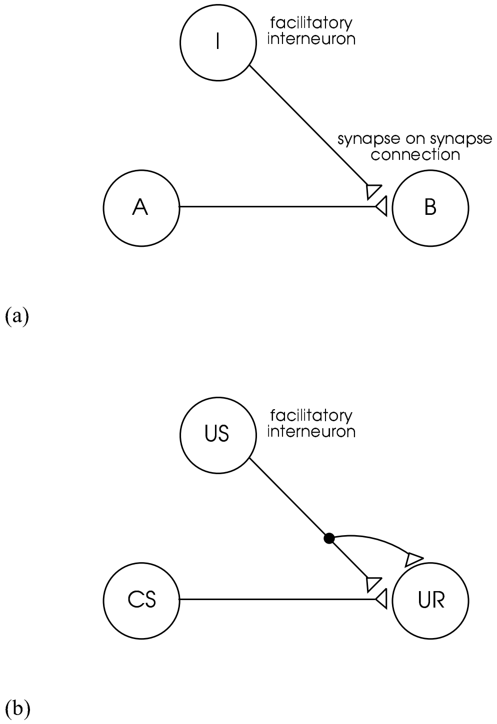


Figure 3: Learning in ARBIB is via mechanisms of (a) sensitisation and (b) classical conditioning modelled by synapse-on-synapse connections.

3.2.3 *Habituating type*

The magnitude of the synaptic weight w_k shown in Figure 2 is simply decreased every time the synapse fires.

$$w_k(t+1) = \begin{cases} w_k(t) - d & \text{if state C} \wedge (w_k(t) > 0) \\ w_k(t) + d & \text{if state C} \wedge (w_k(t) < 0) \\ w_k(t) & \text{otherwise} \end{cases}$$

where d is a constant decrement (typically ~ 1).

3.2.4 *Sensitising type*

The synapse from A to B is strengthened every time I fires, as shown in Figure 3(a), according to:

$$w_{\mathrm{targ}}(t{+}1) = \begin{cases} w_{\mathrm{targ}}(t) + w_{\mathrm{sos}}(t) & \text{if states B, C D} \\ w_{\mathrm{targ}}(t) & \text{otherwise} \end{cases}$$

where ‘targ’ denotes the target synapse (to be sensitised) and ‘sos’ denotes the synapse-on-synapse influence.

3.2.5 *Conditioning type*

As shown in Figure 3(b), the synapse from conditioned stimulus (CS) to unconditioned response (UR) is strengthened proportionally to the interstimulus interval (ISI) between the CS and the US firing:

$$w_{\mathrm{targ}}(t+1) = \begin{cases} w_{\mathrm{targ}}(t) + k w_{\mathrm{sos}}(t) & \text{if } \texttt{fired}_{\mathrm{sos}} \\ w_{\mathrm{targ}}(t) & \text{otherwise} \end{cases}$$

where:

$$k = \frac{nT}{\psi} e^{\left(\frac{-nT}{\varsigma}\right)}$$

and nT is a count of sample periods initiated by encountering state C for the target neuron, ψ (= 250) is an empirically-set scaling factor, ς (= 500) is a constant chosen to maximise the effect of conditioning when the conditioning stimulus precedes the unconditioned stimulus by 0.5 s, as shown in Figure 4.

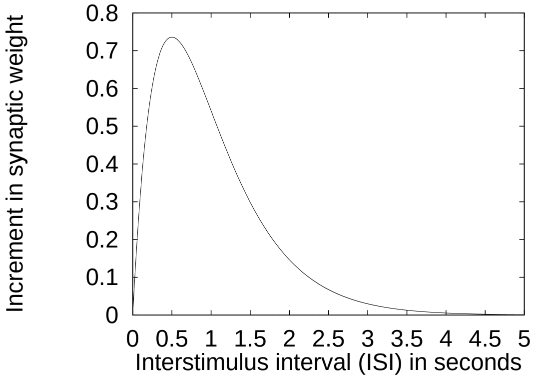


Figure 4: Increment in synaptic weight through conditioning as a function of the ISI.

4. Long-Term Memory

Synaptic plasticity in ARBIB’s nervous system, provided by models of associative classical conditioning and non-associative sensitisation and habituation, allows it to learn collision avoidance skills by switching from short-range to long-range sensory modalities. However, once the long-range modality has become dominant in eliciting the avoidance reflex, this conditioned response will eventually extinguish via autonomous decay (Sutton and Barto 1981; Damper, French, and Scutt, forthcoming), which previously constituted ARBIB’s short-term (and only) memory. Thus, the useful information is lost until it is once again re-acquired through stimulus substitution for a collision avoidance reflex. It therefore needs a way to keep useful information acquired by its short-term memory before it extinguishes.

In this section, we consider the clue that biology has given us for building a long-term memory. Sensitisation experiments with *Aplysia* have linked long-term memory with the growth of new synaptic connections (Bailey and Kandel 1994, p. 42). Studies of rats given reaching tasks have also shown an increase in synapse count (Greenough, Larson, and Withers 1985).

4.1 Synaptogenesis

Synaptogenesis, which has been found to occur after long term potentiation (LTP), duplicates (and so strengthens) localised synapse connectivity. Hence, Toni, Buchs, Nikonenko, Bron, and Muller (1999, p. 421) write:

> "As pharmacological blockade of LTP prevented these morphological changes, we conclude that LTP is associated with the formation of new, mature and probably functional synapses contacting the same pre-synaptic terminal and thereby duplicating activated synapses."

Also, Levitan and Kaczmarek (1997, p. 501) write on the subject of associative LTP:

> "The requirements for temporal pairing of the two stimuli are identical to those required for associative learning paradigms."

4.2 Augmenting the Cell Biological Alphabet

Can a temporal relationship can be used to decide the strength of new synapses? In the spirit of Hawkins and Kandel's cell biological alphabet, the temporal relationship between the unconditioned and conditioned stimuli for classical conditioning has been applied to synaptogenesis in Hi-NOON. Here, the newly-created synapse has a strength calculated from the difference between the elapsed times of post-synaptic cell firing and conditioned-synapse firing.

Pursuing this idea, it was decided to constrain the formation of new synapses by introducing a new predicate into the Hi-NOON model. This is simply the condition that a new synapse is only created, parallel to an existing conditioned synapse, once the conditioned synaptic strength reaches some percentage (say, 90%) of the allowed maximum.

Thus, the cell biological alphabet has been extended in Hi-NOON. First the mechanisms of sensitisation have been applied to classical conditioning (Damper, French, and Scutt, forthcoming), and second, the temporal relationship from classical conditioning has been applied to synaptogenesis.

4.3 Obstacle Avoidance Experiments

A Nomad Scout 2 (see `http://www.robots.com`) controlled through the Nomadic Technologies Nserver software via the Scout's host port was used as the ARBIB platform for all experiments in this paper. The Scout has 6 bump sensors and 16 sonar devices arranged around its circumference as shown in Figure 5.

A total of 14 runs was carried out. Each 11 minute run consisted of ARBIB having free range to travel around the robot laboratory, negotiating obstacles in its path. The first seven runs were made with synaptogenesis enabled in the Hi-NOON model. For comparison, the remaining runs were made with synaptogenesis disabled. By interaction with the environment, ARBIB learns (by stimulus substitution) to elicit its avoidance reflex. Hence, the measure of bump sensory neuron activity (the unconditioned stimulus) is a useful indicator of how successfully it has learned to avoid direct contact with obstacles.

Figure 5: The Nomad Scout 2 robot used in the obstacle avoidance experiments.

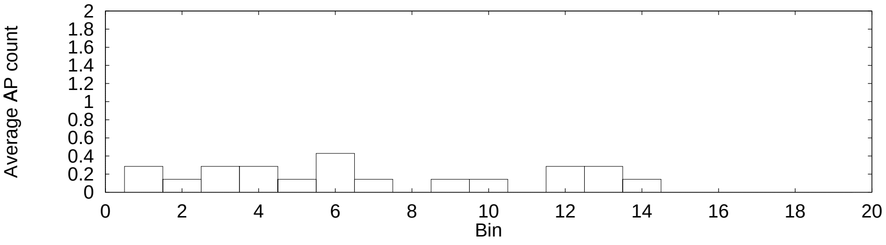
4.4 Results

Figure 6 shows the average action potential count of the left bump sensory neuron for bins of 100 sample points throughout the tests. Results of the first seven runs are shown in Figure 6(a). Comparing these results with Figure 6(b), we see a decrease in activity during the runs with synaptogenesis enabled.

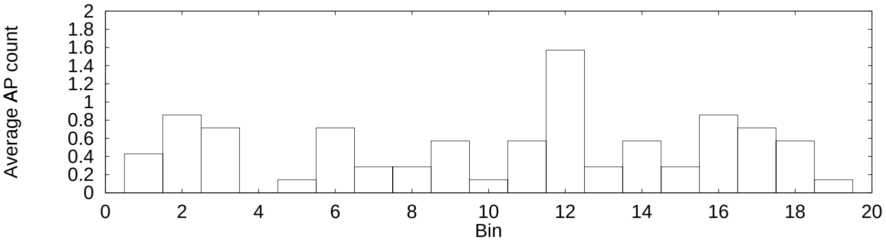
An important observation, obtained by examining log files of synaptogenesis events generated during the experiment, is that ipsilateral sensory and interneurons had undergone greater synaptogenesis than contralateral neurons. This means that sonar sensors had started to act, through stimulus substitution, for bump sensors on the same side (Damper, French, and Scutt, Figure 5, forthcoming). If the sonar sensory transducers were to be swapped over, we expect that contralateral sensory-interneuron pairs would become dominant instead. This experiment remains to be done.

5. Medium-Term Memory

ARBIB cannot associate stimuli which have a large ISI (more than two seconds, say) as this results in only a very small strengthening of a conditioned synapse (Figure 4). Such small changes in synaptic strength make very little differ-



(a) Left bump sensory neuron firing activity with synaptogenesis.



(b) Left bump sensory neuron firing activity without synaptogenesis.

Figure 6: Neural activity in the left bump sensory neurons (a) with and (b) without synaptogenesis.

ence to Arbib's behaviour, and yet in some circumstances being aware of such events could have survival value. An example might be learning the threat display behaviour of another agent deliberately crossing Arbib's path. This is potentially important since an attack might follow some time after the display has occurred. Arbib also needs to be able to ignore constant, neutral background information. Aspects of the environment which hold no survival value but which persist over a large period of time should not consume attention and processing resources. The following two subsections introduce the basis for medium-term memory and habituating sensory processing circuits, inspired from the mechanical 'tortoise' *Machina docilis* of Grey Walter (1951) and the medial pallium model of Wang and Arbib (1992), respectively.

5.1 Machina docilis

Machina docilis (Grey Walter 1951) can associate a neutral stimulus with an appetitive or aversive one depending upon the internal configuration of the circuit. Its memory is stored as the decaying feedback around a valve oscillator, representing a recurrent neural circuit. Grey Walter's schematic has been modified and incorporated into Arbib as `M1`, `M2` and `M3` in Figure 7. Hence, firing `M1`, say, will initiate oscillations. As in the valve circuit, these will decay because of an habituating synapse, `D` between `M1` and `M2`. Hence, the time of oscillation is governed by the habituation rate. The decaying oscillations through `M1`, `M2` and `M3` form the medium-term memory.

5.2 Medial Pallium Model

The medial pallium (MP) model of Wang and Arbib (1992) is one part of a larger model which simulates some of the circuitry for orienting and prey catching behaviour in the toad. The MP receives input from the toad's retina and anterior thalamus (AT), giving an output which affects the tectum, where prey catching behaviour is generated.

We have included a small circuit inspired from Wang and Arbib's medial pallium model as shown in Figure 7. Instead of this simplified MP model receiving input from models of a retina and anterior thalamus, it comes from the activity of a sonar (range-finding) sensory cell, which thereby acts as a simple AT. The closer the target and the greater the depolarisation, the higher the firing rate of the sensory cell and, hence, of the MP circuit. Figure 8(a) shows the output from the sensory cell. This circuit gives Arbib the ability to ignore constant stimuli through habituation, and yet dishabituate to a change in stimulus strength.

The `P1` neurons form a simple position-threshold-slope map (Grossberg and Kuperstein 1986) that converts different input intensities to different positions of firing activity in an array of neurons. Here, with an array of just two `P1` cells, Arbib can differentiate between high and low firing rates of the sensory cell. There is a hierarchy of `MP2` cells in which the `P1 HIGH` cell inhibits the `MP2 LOW` cell. The `MP2 HIGH` cell connects to `MP3` and `P2` cells, whereas the `MP2 LOW` connects only with its `P2` cell. These connections with `MP3` and `P2` cells are habituating types, and are labelled `A`, `B` and `C` in Figure 7.

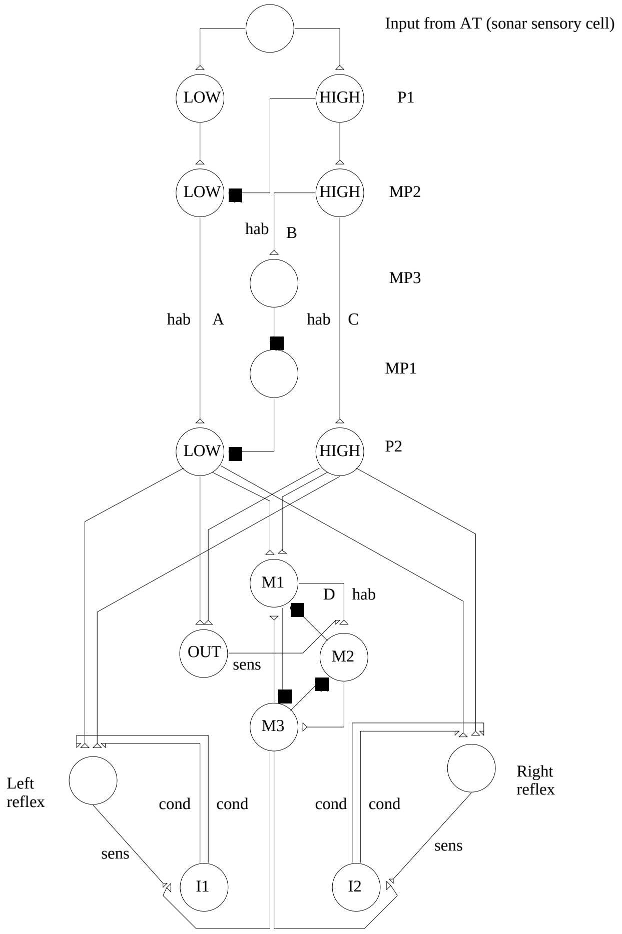


Figure 7: The medial pallium model feeding an oscillatory memory, with a possible interface to ARBIB's reflex interneurons.

5.3 *Results*

Suppose ARBIB is placed into a world where it senses only a distant object, then the sensory cell will have a low firing rate which is insufficient to depolarise `MP2 HIGH` through threshold. Hence, only `MP2 LOW` fires, causing the habituating synapse `A` to fire and firing `P2 LOW`. If another agent now crosses ARBIB's 'line of sight' with the distant object, then the sensory cell will fire at an increased rate (Figure 8(a), samples [0: 3000]), `MP2 HIGH` will fire and `MP2 LOW` will be inhibited. Hence, `P2 HIGH` will start firing (Figure 8(c), [0: 3000]).

Since synapses `A, B` and `C` are habituating types. Their strength will be reduced each time they fire. Hence, with a constant view of the distant object, `A` will habituate to a point where `P2 LOW` will not reach threshold (Figures 8(b) and 8(f) [3000:6000]). Thus, the distant object will be ignored. If our second agent once again crosses ARBIB's path (Figure 8(a), [6000: 9000]), the `P2 HIGH` cell will once again fire because synapse `C` has recovered from its earlier activity. Although the distant object is being ignored, presence of this new and closer stimulus is receiving attention. As both `P2` cells stimulate the cell labelled `out` in Figure 7, we see that its activity habituates to the distant object (Figure 8(d), [3000: 6000]). However, it then dishabituates in the presence of the closer stimulus (Figure 8(d), [6000: 9000]).

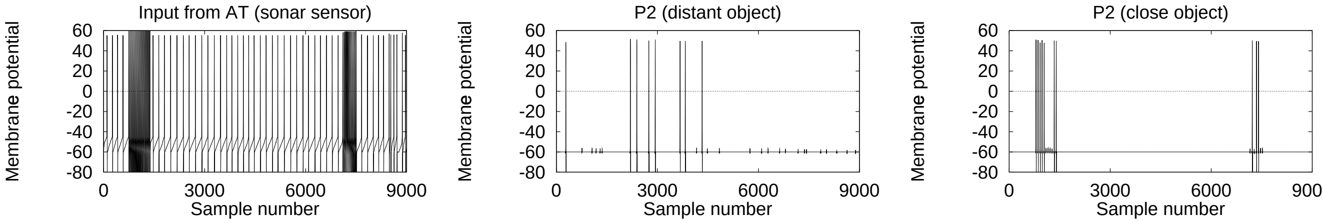
The `P2` cells excite the `M1` cell in the oscillatory memory. Once ARBIB has habituated to the distant stimulus, and is dishabituated to the transitory close stimulus, the oscillatory memory will store this event as an habituating activity around its loop. This is shown in Figure 8(e) [6000: 9000].

Cell `M3` has only weak excitatory connections with interneurons `I1` and `I2`. The right and left reflex interneurons make sensitising synapse-on-synapse connections with the `M3`, `I1` and `I2` synapses, with the `I1` and `I2` cells making conditioning synapse-on-synapse connections with `P2`-to-reflex synapses. Thus, if ARBIB dishabituates to our second agent, which some time later causes it to experience an aversive stimulus and so fire the reflex interneurons, the connections between the `P2 HIGH` cell and the appropriate reflex interneuron will be strengthened. This will occur for the duration of the oscillatory memory, which acts to amplify the weak conditioning – because of a potentially large ISI between the `P2` activity and the reflex activity.

6. Discussion

The main purpose of this work has been to provide stability of learning in the behaving animat ARBIB. Our philosophy throughout has been to base learning, and thereby behaviour, on plausible biological principles. The question therefore arises: can ARBIB do anything that other robots or animats explicitly programmed (e.g. using reinforcement learning) to perform low-level tasks cannot do? Some examples of the latter approach are seen in the publications of Verschure, Kröse, and Pfeifer (1992), Zalama, Gaudiano, and Coronado (1995) and Gaudiano, Guenter, and Zalama (1997) (although these have not always used a real robot platform). The answer to the above question is that there is probably very little if any difference in the observed behaviours. But this is perhaps to miss the point. Because we have modelled ARBIB very firmly on known or inferred neural mechanisms, we can trace a clear link between physiology and behaviour. In particular, in this paper, we have shown how a simplified version of synaptogenesis (currently receiving great attention in experimental neurobiology) can neatly solve the well-known stability-plasticity dilemma of learning in a real-world robot.

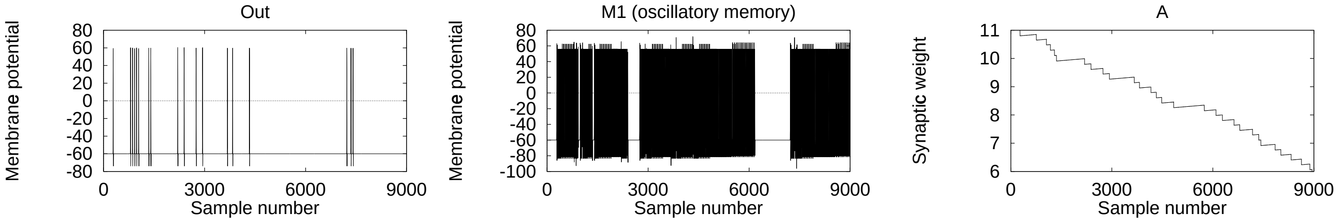
It is, of course, always possible to use engineering, or non-biological, mechanisms to achieve desired robot behaviour. Examples of this are the work of Nehmzow, Smithers, and McGonigle (1995) who use a combination of (PDP-style) neural networks and high-level 'instinct' rules to increase behavioural repertoire in a mobile robot, and Krebs and Bossel (1997) who use symbolic classifier systems as the basis of evolved behaviour. While such work is valid and useful in its own terms, it does not illuminate links between neurophysiology and behaviour – such as the way that symbolic capacity might be grounded in robotic capacity (Harnad 1990, 1995) –



(a) Activity in the sonar sensory cell. Note the high frequency bursts and compare with Figure (c).

(b) Activity in the `P2 LOW` cell. Compare its habituation with Figure (f).

(c) `P2 HIGH` cell activity. Compare with Figure (a).



(d) Activity in the `OUT` cell. Compare with Figures (b) and (d).

(e) Activity in the medium-term memory cell `M1`. Compare with Figure (d).

(f) The weight of habituating synapse `A`.

Figure 8: Activity in: (a) `Sonar sensory`, (b) `P2 LOW`, (c) `P2 HIGH`, (d) medial pallium `OUT`, and (e) `M1` memory cells, with (f) the weight of synapse `A`.

in the way that we have tried to do here.

7. Conclusions and Future Work

This paper has reported experiments incorporating a simplified form of synaptogenesis as the basis of a stable long-term memory mechanism in the ARBIB mobile robot. It has also examined an idea for a medium-term memory, based upon the work of Grey Walter (1951), which in turn is fed from a circuit derived from Wang and Arbib's (1992) model of the medial pallium. This gives the robot the beginnings of an ability to ignore neutral stimuli in its environment, while also attempting to overcome the problem of conditioning with stimuli which have a large inter-stimulus interval. Thus, ARBIB's competence is evolving as its nervous system expands. This will continue as more advanced robot platforms utilising colour vision and manipulator capabilities are employed, allowing greater interaction with the world and sustaining growth of the nervous system.

Acknowledgements

Many thanks to Ulrich Nehmzow, Ten Min Lee, Stephen Marsland, Jason Fleischer and Andrew Pickering of the Robot Research Laboratory, Department of Computer Science, University of Manchester, for their interest, help and use of equipment. The work reported here was supported by a studentship from the UK Engineering and Physical Sciences Research Council.

References

Bailey, C. H. and E. R. Kandel (1994). Structural changes underlying long-term memory storage in *Aplysia*: A molecular perspective. *The Neurosciences 6*, 35–44.

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception 1*, 377–394.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence 47*, 139–159.

Carpenter, G. A. and S. Grossberg (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *IEEE Computer 21*(3), 77–88.

Churchland, P. M. (1986). Some reductive strategies in cognitive neurobiology. *Mind 95*(379), 279–309.

Clark, A. (1987). Being there: Why implementation matters to cognitive science. *Artificial Intelligence Review 1*, 231–244.

Coad, P. and E. Yourdon (1991). *Object Oriented Analysis*. Englewood Cliffs, NJ: Prentice-Hall. Second Edition.

Damper, R. I., R. L. B. French, and T. W. Scutt (forthcoming). ARBIB: an autonomous robot based on inspirations from biology. *Robotics and Autonomous Systems*, in press.

Gaudiano, P., F. H. Guenter, and E. Zalama (1997). The neural dynamics approach to sensory-motor control: Overview and recent applications in mobile robot control and speech production. In O. Omidvar and P. van der Smagt (Eds.), *Neural Systems for Robotics*, pp. 153–194. San Diego, CA: Academic.

Greenough, W. T., J. R. Larson, and G. S. Withers (1985). Effects of unilateral and bilateral training in a reaching task on dendritic branching of neurons in the rat motor sensory forelimb cortex. *Behavioral and Neural Biology 44*, 301–314.

Grey Walter, W. (1951). A machine that learns. *Scientific American 185*(5), 60–63.

Grossberg, S. and M. Kuperstein (1986). *Neural Dynamics of Adaptive Sensory-Motor Control*. Amsterdam: Elsevier (North-Holland).

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

Harnad, S. (1995). Grounding symbolic capacity in robotic capacity. In L. Steels and R. Brooks (Eds.), *The Artificial Life Route to Artificial Intelligence: Building Situated Embodied Agents*, pp. 277–286. New Haven: Lawrence Erlbaum.

Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. Cambridge, MA: Bradford Books/MIT Press.

Hawkins, R. D. and E. R. Kandel (1984). Is there a cell biological alphabet for simple forms of learning? *Psychological Review 91*, 375–391.

Krebs, F. and H. Bossel (1997). Emergent value orientation in self-organization of an animat. *Ecological Modelling 96*, 143–164.

Levitan, I. B. and L. K. Kaczmarek (1997). *The Neuron, Cell and Molecular Biology*. New York, NY: Oxford University Press.

MacGregor, R. J. (1987). *Neural and Brain Modeling*. London, UK: Academic.

Nehmzow, U., T. Smithers, and B. McGonigle (1995). Increasing behavioural repertoire in a mobile robot. In J.-A. Meyer, H. Roitblat, and S. W. Wilson (Eds.), *From Animals to Animats 2: Proceedings of the 2nd International Conference on Simulation of Adaptive Behavior*, pp. 291–297. Cambridge, MA: Bradford Books/MIT Press.

Scutt, T. W. and R. I. Damper (1991). Computational modelling of learning and behaviour in small neuronal systems. In *Proceedings of International Joint Conference on Neural Networks,* Singapore, pp. 430–435.

Scutt, T. W. and R. I. Damper (1997). Designing a nervous system for an adaptive mobile robot. In A. Browne (Ed.), *Neural Network Perspectives on Cognition and Adaptive Robotics*, pp. 220–250. Bristol, UK: Institute of Physics Press.

Selverston, A. I. (1993). Modeling of neural circuits – What have we learned? *Annual Review of Neuroscience 16*, 531–546.

Sutton, R. S. and A. G. Barto (1981). Towards a modern theory of adaptive networks: Expectation and prediction. *Psychological Review 88*, 135–170.

Toni, N., P. A. Buchs, I. Nikonenko, C. R. Bron, and D. Muller (1999). LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite. *Nature 402*, 421–425.

Verschure, P. F. M. J., B. J. A. Kröse, and R. Pfeifer (1992). Distributed adaptive control: The self-organization of structured behavior. *Robotics and Autonomous Systems 9*, 181–196.

Wang, D. and M. A. Arbib (1992). Modeling the dishabituation hierarchy: The role of the primordial hippocampus. *Biological Cybernetics 67*, 535–544.

Wilson, S. W. (1991). The animat path to AI. In J.-A. Meyer and S. W. Wilson (Eds.), *From Animals to Animats: Proceedings of the 1st International Conference on Simulation of Adaptive Behavior*, pp. 15–21. Cambridge, MA: Bradford Books/MIT Press.

Zalama, E., P. Gaudiano, and J. L. Coronado (1995). A real-time, unsupervised neural network for the low-level control of a mobile robot in a nonstationary environment. *Neural Networks 8*(1), 103–123.