



www.sciencemag.org/cgi/content/full/318/5853/1103/DC1

Supporting Online Material for
Learning in and from Brain-Based Devices

Gerald M. Edelman

E-mail: edelman@nsi.edu

Published 16 November 2007, *Science* **318**, 1103 (2007)
DOI: 10.1126/science.1148677

This PDF file includes:

SOM Text
Figs. S1 to S4
References

Supplementary Online Material

Learning in and from Brain-Based Devices

Gerald M. Edelman, M.D., Ph.D.
The Neurosciences Institute
San Diego, California
edelman@nsi.edu

Brain-based device (BBD) behavior is guided by a simulated nervous system modeled on the anatomy and physiology of the mammalian nervous system but, obviously, with far fewer neurons. The simulation consists of a number of areas labeled according to the analogous neocortical, hippocampal and subcortical brain regions. Each area contains neuronal units that can be either excitatory or inhibitory, and each neuronal unit represents a local field potential generated by a population of approximately 100 neurons firing over a period of roughly 200 milliseconds. To distinguish modeled areas from corresponding regions in the mammalian nervous system, the simulated areas are indicated in italics (e.g., *IT*).

During each simulation cycle sensory input is processed, the states of all neuronal units are computed, the connection strengths of all plastic connections are determined, and motor output is generated. In our experiments, execution of each simulation cycle required about 200 ms of real time. The neural simulation was run on a Beowulf cluster of up to 12 Pentium IV computers running the Linux operating system. All sensory input from the brain-based device and motor commands to the device were communicated through wireless links between the device and one of cluster's workstations. During each simulation cycle, all neuronal activities were saved on a hard disk, and behavioral measurements were recorded.

This supplemental text covers two very different models, Darwin VII (investigating the acquisition of conditioned responses) and Darwin X (investigating spatial memory). Darwin VII's simulation included models of visual, auditory, inferotemporal, and motor cortices along with the value system and the colliculus. Darwin VII had 20,000 neuronal units in 18 areas and a half million synaptic connections. Darwin X's simulation included models of visual, inferotemporal, parietal, and motor cortices, along with the value system, basal forebrain, anterior thalamic nucleus and the medial temporal lobe. Darwin X had 100,000 neuronal units in 50 areas and one and a half million synaptic connections.

The following two sections describe the particular behavioral and neuroanatomical traits of Darwin VII and Darwin X. The final section addresses the detailed algorithms for neural simulation which underly both models, but omits the long tables that describe each of these models in great detail. Those who are interested in these parameters can find them in the original publications (1–3).

1 Darwin VII

1.1 Behavior

The acquisition of conditioned responses in Darwin VII was assessed through a task that required differentiating between different types of visual or auditory patterns in order to obtain appetitive and avoid aversive “taste” stimuli. Darwin VII operated in an enclosed area with black cloth-covered walls and a floor covered with opaque black plastic panels, on which we distributed stimulus blocks (6 cm metallic cubes; Figure 1). The top surfaces of these blocks were covered with removable black and white patterns. All other surfaces of the cubes were featureless and black. All experiments reported in this paper were carried out with block stimulus exemplars of two basic designs: blobs (several white patches 2-3 cm in diameter) and stripes (width 0.6 cm, evenly spaced). Stripes, when picked up with Darwin VII’s gripper, could be viewed either in a horizontal or vertical orientation, yielding a total of three stimulus classes (blob, horizontal and vertical) of visual patterns to be discriminated. A flashlight aligned with the CCD camera, caused the blocks, which contained photodetectors, to emit a tone when Darwin VII was in the general area. The sides of the stimulus blocks were metallic and could be rendered either strongly or weakly conductive. In the experiments described in this paper, negative value blocks had weakly conductive “taste” with a blob visual pattern and a 3.3 kHz tone, whereas positive value blocks had strongly conductive “taste” with a striped visual pattern and a 3.9 kHz tone.



Figure 1: Darwin VII in its environment.

Darwin VII's body consisted of a mobile robotic base (16 inches in diameter) equipped with several sensors and effectors, and was capable of communicating wirelessly with a neural simulation running on a remote computer. The device had wheels that permitted independent translational and rotational motion, pan/tilt movement for its camera and microphones, and object gripping with a one degree-of-freedom manipulator (see Figure 1).

Darwin VII had three built-in behaviors. It moved forward by default. If it encountered the boundaries of the environment (detected by eight infrared (IR) sensors mounted at 45-degree intervals around the mobile platform) it backed away. If a block entered the gripper (detected by IR sensors in the gripper) then the gripper closed, the object was picked up, and conductivity was measured across its exposed contacts.

The neural simulation activity (described below) determined whether Darwin VII would approach or back away from a block once one entered the camera's visual field. Once a block was approached and in the gripper Darwin VII would "taste" the block's conductivity. The "taste" of the block triggered activity in the value system of the simulated nervous system (see Action selection subsection, below). Initially Darwin VII picked up blocks of positive and negative value equally and randomly, but after experience and the changes induced in synapses by value system activity Darwin X learned to discriminate the blocks based on visual and auditory cues and reliably made the appropriate responses.

Training was repeated with seven different Darwin VII subjects. Each subject consisted of the same physical device and gross anatomical structure, but each possessed a unique micro-circuitry as a consequence of random initialization of the simulation and each had a separate history of interaction with the environment. The group of Darwin VII subjects demonstrated greater than 90% correct responses to positive and negative value blocks after training for 10 presentations of positive value blocks and 10 presentations of negative value blocks.

1.2 Sensory input

Visual images from the camera were clipped, such that only the center square of the image remained, and spatially averaged to produce a 64x64 pixel image with each pixel normalized between 0 (black) and 1 (white), and mapped directly to neuronal units of area R (analogous to retinal output) in the neural simulation. For visual tracking, the image was mapped directly to area C in a similar fashion as area R , but without clipping, such that C had a wide-angle view (see Figure 2).

The activity of neuronal units in the primary visual areas, VA_p , were selective for blob-like features, short horizontal, or vertical line segments. Activity of neurons in VA_p were calculated by performing a 2-dimensional convolution on area R , the convolution kernel differing according to whether a particular neuronal unit was sensitive to horizontal lines, vertical lines, or blobs. Neuronal units within VA_p were retinotopic, i.e. topographically mapped to the visual image, and projected non-topographically via activity-dependent plastic connections to a secondary visual area, analogous to the inferotemporal cortex (IT). IT contained local excitatory and inhibitory interactions producing activity patterns that were characterized by focal regions of excitation surrounded by inhibition.

The frequency and amplitude information captured by Darwin VII's microphones was relayed to a simulated cochlea area ($LCoch$, $RCoch$). $LCoch$ and $RCoch$ each had 64 neuronal units and each cochlear neuronal unit had a cosine tuning curve with a tuning width of 1 kHz and a preferred frequency which over the ensemble of units ranged from 2.9 kHz to 4.2 kHz. Activity of a cochlear neuronal unit was obtained by multiplying the value from the cosine tuning curve with the amplitude of the microphone

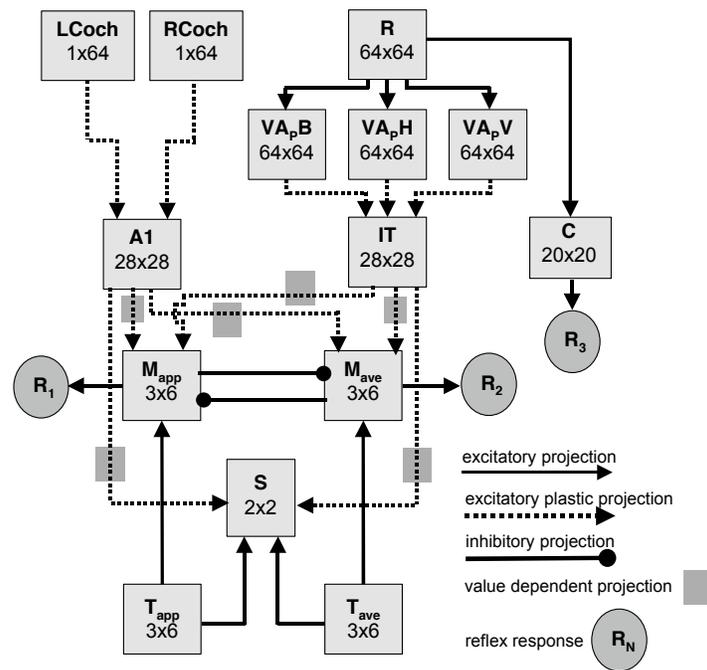


Figure 2: Schematic of the anatomy of Darwin VII's nervous system. The neural areas and the number of neuronal units in each neural area are given in the shaded boxes in the figure. R1 and R2 correspond to appetitive and aversive behavioral responses, and R3 corresponds to a visual tracking response. Inhibitory connections are omitted for clarity.

signal. Cochlear neuronal units projected tonotopically (i.e., were frequency mapped) to neuronal units in neural area *AI*. Similar to area *IT*, *AI* contained local excitatory and inhibitory interactions.

1.3 Action selection

Two neuronal areas were capable of triggering appetitive (*Mapp*) or aversive (*Mave*) behaviors. The behavioral responses were triggered if the difference in instantaneous activity between motor areas *Mapp* and *Mave* exceeded a behavioral threshold ($\beta = 0.3$). The taste system consisted of two kinds of sensory units responsive to either strong conductivity (*Tapp*) or weak conductivity (*Tave*). The taste system sent non-plastic projections to the motor areas (*Mapp*, *Mave*) and the value system (*S*). *AI* and *IT* sent plastic projections to the motor areas (*Mapp*, *Mave*) and the value system (*S*). Initially, only the taste system was strong enough to elicit a motor behavior. Area *S* projected diffusely, with long-lasting value-dependent activity to the auditory, visual and motor behavior neurons. The visual tracking system, which controlled the approach to objects identified by brightness contrast with respect to the background, was achieved by connections from the retinal area *R* to area *C* (colliculus), which in turn sent motor commands to the device’s wheels. The connection strengths to achieve tracking behavior were fixed based on learning experiments in a previous study.

Activation of the value system (see Area *S* and value dependent projects in Figure 2) signaled the occurrence of salient sensory events and contributed to the modulation of connection strengths of all synapses in the affected pathways. For example, tasting a block is a salient event with the consequence that, at that time, behavior should be reinforced or weakened through synaptic change. Area *S* thus is analogous to an ascending neuromodulatory system in that its units show uniform phasic responses when activated and its output acts diffusely over multiple pathways by modulating synaptic change (4).

Value, V , was calculated at every time step of the simulation according to

$$V(d) = 1 + f(d) \frac{\bar{S} + V(d-1) \times (d-1)}{d} \quad (1)$$

where d is the delay of the value term and is incremented every simulation cycle after the onset of area *S* activity with a range of 1 through 9, \bar{S} is the average activity in area *S* at this time point, and $V(d-1)$ is the value of V at time $d-1$. f is a convolution function that scales the activity over the delay period with values of 0.1, 0.1, 0.3, 0.7, 1.0, 1.0, 0.7, 0.3 and 0.1 for the 9 delay increments. The effect of this convolution is to delay onset of the value system activity and spread the activity over time.

2 Darwin X

2.1 Behavior

Spatial memory in Darwin X was assessed in a dry variant of the Morris water maze task (5) in which the BBD is rewarded by finding a hidden platform. Successful performance of this task is reflected by the BBD navigating to the hidden platform from any starting position by using only visual landmarks and self-movement cues. Darwin X was allowed to explore an enclosure in which there were visual landmarks on the walls and a circular “hidden platform” of reflective black paper (see Figure 3). The platform could not be detected by the visual system of Darwin X, but was detectable at close range by an infrared (IR) sensor on the front of the device. This IR sensor triggered activity in the value system

of the simulated nervous system (see Action selection subsection, below). Initially Darwin X found the platform through random exploration of the environment but, after experience and the changes induced in synapses by value system activity, Darwin X learned to repeatably and directly approach the hidden platform.

Darwin X began a trial from one of four starting locations (see Figure 3 A) and explored the enclosure until it encountered the hidden platform or until a time limit of 1,000 s was reached. A training block was defined as a set of four trials from each of four starting locations. Four blocks (16 trials) were completed by the device during training. Training was repeated with nine different Darwin X subjects. Each subject consisted of the same physical device and gross anatomical structure, but each possessed a unique micro-circuitry as a consequence of random initialization in the simulation and each had a separate history of interaction with the environment.

The group of Darwin X subjects showed significant improvement in the hidden platform task, as measured by the time to find the hidden platform (search time), as training progressed. The median search times of the last four trials were significantly shorter than the first four trials (223.5 s in trials 1316 and 532.1s in trials 14, $p < 0.01$, Wilcoxon sign rank test). In general, after the second block (trials 916), Darwin X traversed directly to the hidden platform from multiple starting points.

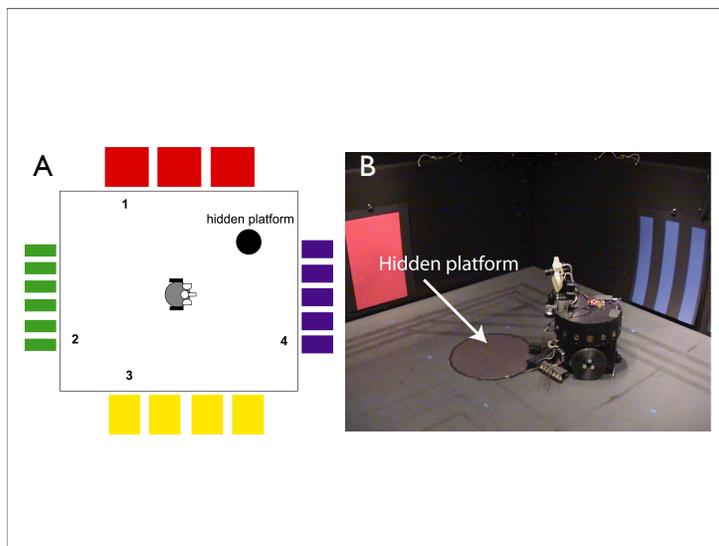


Figure 3: Layout of the enclosure used for the hidden platform task. **A:** Schematic of the environment. Enclosure is 16 feet by 14 feet with black walls and flooring. Pieces of differently colored paper of varying widths were hung on each of the walls. A hidden circular platform, 24 inches in diameter and made of reflective black paper, was placed in the center of a quadrant in the enclosure. Each trial began in one of four starting locations (see numbers 1-4 in the diagram). **B:** Snapshot of Darwin X in its environment.

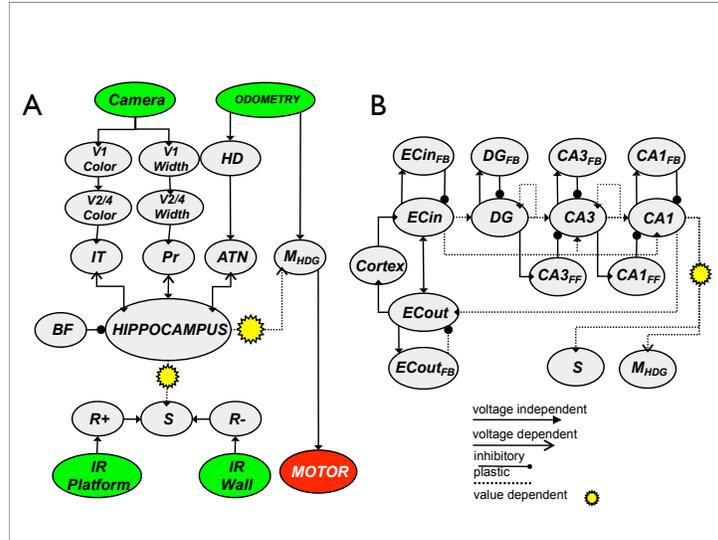


Figure 4: Schematic of the regional and functional neuroanatomy of Darwin X. Gray ellipses denote different neural areas. Arrows denote projections from one area to another. **A:** Diagram of cortical-hippocampal connectivity. Input to the neural simulation comes from a camera, wheel odometry, and IR sensors for wall and platform detection. The simulation contains neural areas analogous to visual cortex (*V1*, *V2*), inferotemporal cortex (*IT*), parietal cortex (*Pr*), head direction units (*HD*), anterior thalamic nuclei (*ATN*), motor areas for egocentric heading (M_{HDG}), a value system (*S*), and positive and negative reward areas (textitR+, R-). The hippocampus is connected with the three major sensor input streams (*IT*, *Pr*, *ATN*), the motor system (M_{HDG}), and the value system (*S*). The hippocampus receives rhythmic inhibition from a simulated basal forebrain region (*BF*). **B:** Diagram of connectivity within the hippocampal region. The modeled hippocampus contains areas analogous to entorhinal cortex (*ECIN*, *ECOUT*), dentate gyrus (*DG*), and the CA3 and CA1 subfields (*CA3*, *CA1*). These areas contain interneurons that implement feedback inhibition (e.g., $CA3 \rightarrow CA3_{FB} \rightarrow CA3$) and feed-forward inhibition (e.g., $DG \rightarrow CA3_{FF} \rightarrow CA3$)

2.2 Sensory input

Vision

Visual images from Darwin X's CCD camera were filtered for color and edges. The filtered output directly affected neural activity in area *VI*, which is composed of functionally segregated sub-areas for color and shape. The CCD camera produced an 80x60 pixel RGB image. Different sized Gabor filters (2x2, 4x14, 16x16, and 32x32) were used to detect vertical edges of varying widths. The output of the Gabor function mapped directly onto the neuronal units of the corresponding *VI* sub-area (*VI-width2*, *VI-width4*, *VI-width16*, and *VI-width32*). Different color filters (red, green, yellow, and blue) were applied to the image. The outputs of the color filters were mapped directly onto the neuronal units of *VI-red*, *VI-green*, *VI-blue*, and *VI-yellow*. *VI* color neuronal units projected to corresponding *V2/4* units and then non-topologically to inferotemporal cortex *IT*. *VI* edge units projected to corresponding *V2/4* units and then retinotopically to parietal cortex *Pr*.

Head direction

A head direction system was modeled after areas of the rodent nervous system (e.g. anterior thalamic nuclei) that respond selectively to the animal's heading (6, 7). Neurons in these areas are often called head direction cells. Odometer information obtained from Darwin X's wheels was used to estimate current heading. This information was input into the head direction neural area (*HD*). Each of the 360 *HD* neuronal units had a cosine tuning curve, which responded maximally to a preferred heading with a tuning width of π radians:

$$(\cos(HD_i - curr_heading))^5 \quad (2)$$

where HD_i is a head direction cell with a preferred direction of $(\frac{i}{360}2\pi)$ and i ranges from 0 to 359.

The head direction cells projected topographically to an area analogous to the anterior thalamic nucleus ($HD \rightarrow ATN$) and to a motor area ($HD \rightarrow M_{HDG}$) used for selecting a new heading (see below).

2.3 Hippocampus

The architecture of the simulated hippocampal formation was based on the known literature about the gross connectivity and micro-architecture of rodent neuroanatomy. This approach is distinct from other robotic navigation models of hippocampal function (8–10) in that it focusses on the how a large-scale model of this anatomy can produce known single-unit phenomena with experience-dependent plasticity. In this model, the anatomical connectivity is fixed, as are synaptic strengths in the sensory input streams, but there are plastic connections within the hippocampus and between the hippocampus and the neural areas responsible for action selection.

The input streams into the hippocampus are from the input areas of the simulation ($ATN \rightarrow EC_{IN}$, $IT \rightarrow EC_{IN}$, $PR \rightarrow EC_{IN}$). Parameter values for the neuronal units and connections in these areas were tuned such that each cortical area (*ATN*, *PR*, and *IT*) had an equivalent synaptic influence on *EC_{IN}*. The relative numbers of neuronal units in each area, and the intrinsic and extrinsic of connectivity of the hippocampus were implemented based on known anatomical measurements (11–13). The perforant path projects mainly from entorhinal cortex to the dentate gyrus but also to the CA3 and CA1 sub-fields ($EC_{IN} \rightarrow DG$, $EC_{IN} \rightarrow CA3$, $EC_{IN} \rightarrow CA1$). The mossy fibers ($DG \rightarrow CA3$), Schaffer collaterals ($CA3 \rightarrow CA1$), and divergent projections from the hippocampus back to cortex ($CA1 \rightarrow EC_{OUT} \rightarrow ATN, IT, PR$)

were also reflected in the neural simulation. Moreover, the prevalent local reentrant connectivity found in the hippocampal formation was included in the model ($EC_{IN} \rightarrow EC_{OUT}$, $DG \rightarrow DG$, and $CA3 \rightarrow CA3$).

There are distinct patterns of intrinsic and extrinsic, feedback and feedforward inhibitory connections in the hippocampal circuitry (12, 14). Feedback inhibitory connections ($EC \rightarrow EC_{FB} \rightarrow EC$, $DG \rightarrow DG_{FB} \rightarrow DG$, $CA3 \rightarrow CA3_{FB} \rightarrow CA3$, and $CAI \rightarrow CAI_{FB} \rightarrow CAI$) and feedforward inhibitory connections ($EC \rightarrow DG_{FF} \rightarrow DG$, $DG \rightarrow CA3_{FF} \rightarrow CA3$, and $CA3 \rightarrow CAI_{FF} \rightarrow CAI$) were included in the model. These connections were important for separating inputs and maintaining network stability.

A simplified model of the basal forebrain provided an extrinsic theta rhythm for the neural simulation. The function of the simulated basal forebrain area was to gate input into the hippocampus and keep activity levels stable. The *BF* area had a rhythmic activity over 13 simulation cycles:

$$BF(t) = \text{theta}(t \bmod 13) \quad (3)$$

where $\text{theta} = \{0.01, 0.165, 0.33, 0.495, 0.66, 0.825, 1.00, 0.825, 0.66, 0.495, 0.33, 0.165, 0.01\}$. *BF* projected to all hippocampal areas with inhibitory connections ($BF \rightarrow EC_{IN}$, EC_{OUT} , DG , $CA3$, CAI). This theta rhythm was spread over 13 simulation cycles in order to ensure that the smoothness of the discretized sine wave and also to allow sufficient time for new input to propagate through the network. Owing to computational limitations, the duration of the cycle was not comparable to that of real nervous systems. The level of inhibition, which was adaptive, kept the activity in hippocampal regions within specific ranges:

$$\begin{aligned} \Delta s f_r(t) &= (s_r(t) - t g t_r) \\ B F_r(t) &= B F(t) + s f_r(t) \end{aligned} \quad (4)$$

where r denotes the region (i.e. EC_{IN} , EC_{OUT} , DG , $CA3$, CAI), $s f_r(t)$ is the scale factor at time t , $s_r(t)$ is the percentage of active neuronal units in region r at time t , $t g t_r$ is the desired percentage of active units in area r ($EC_{IN}=10\%$, $EC_{OUT}=10\%$, $DG=20\%$, $CA3=5\%$, and $CAI=10\%$), and $B F_r(t)$ is the pre-synaptic neuronal unit activity for a connection from *BF* to hippocampus region r .

2.4 Action selection

Activity in the simulated value system (Area *S*, Figure 4) signals the occurrence of salient sensory events and this activity contributes to the modulation of value-dependent connection strengths in synaptic pathways ($CAI \rightarrow S$ and $CAI \rightarrow M_{HDG}$). The projection from our simulated *CAI* to the value and goal decision areas is consistent with the connectivity between *CAI* and nucleus accumbens and frontal areas (15, 16). Initially, *S* is activated by the hidden platform IR detector ($R^+ \rightarrow S$), causing potentiation of value-dependent connections. *S* is also activated by collision with the arena walls ($R^- \rightarrow S$), causing depression of value-dependent connections. After experience, the value system could be activated by *CAI*. The magnitude of potentiation or depression is based on a neural implementation of a temporal difference (TD) learning rule (17, 18). The TD rule applied in this model is:

$$TD(t) = \begin{cases} R^+(t) - \overline{S(t-\tau)}, & R^+ > 0 \\ \overline{S(t-\tau)} - R^-(t), & R^- > 0 \\ S(t) - \overline{S(t-\tau)}, & \text{otherwise,} \end{cases} \quad (5)$$

where $\overline{S(t)}$ is the average activity of the value system at time t , τ is one theta cycle (13 simulation cycles), R^+ is positive reward and equal to 1 if the BBD is over the hidden platform. The basic idea of the temporal difference rule is that learning is based on the difference between temporally successive predictions of rewards. The goal of learning is to make the learner's current prediction of expected reward match more closely the actual expected reward at the next time interval (τ). If the expected reward value increases over τ , TD is positive and affected synaptic connections are potentiated, and if the change in value decreases, TD is negative and affected synaptic connections are depressed. Further details on how the temporal difference is applied to individual synaptic connections are given in the *Neuronal Dynamics* section below.

Darwin X selected a new heading every three theta cycles (39 simulation cycles). The device stopped moving forward, turned 60° counterclockwise, waited for three seconds, then turned clockwise 60° , waited for three seconds, and finally turned clockwise another 60° , and waited three seconds. The average activity of M_{HDG} on each heading was calculated during the wait periods. A softmax algorithm was used to create a probability distribution for choosing a new heading:

$$p(newhdg) = \frac{\exp(40\overline{M_{HDG}(newhdg)})}{\sum_{h \in \{hdg-90, hdg+90\}} \exp(40\overline{M_{HDG}(h)})} \quad (6)$$

where $newhdg$ is a possible new heading for Darwin X, $\overline{M_{HDG}(newhdg)}$ is the average activity of M_{HDG} at a possible new heading, hdg is the current heading, and h has three positions (current heading, current less 60° , and current plus 60°).

3 Neuronal Dynamics and Synaptic Plasticity

Here we describe the algorithmic details of the neural simulation that are common to both Darwin VII and Darwin X.

A neuronal unit is simulated by a mean firing rate model, in which the mean firing rate variable of each unit corresponds to the average activity of a group of roughly 100 real neurons during a time period of approximately 200 milliseconds. Synaptic connections between neural units, both within and between neuronal areas, are set to be either voltage-independent or voltage-dependent, and either plastic or non-plastic. Voltage-independent connections provide synaptic input regardless of post-synaptic state. Voltage-dependent connections represent the contribution of receptor types (e.g. NMDA receptors) that require post-synaptic depolarization to be activated.

The mean firing rate (s) of each neuronal unit ranges continuously from 0 (quiescent) to 1 (maximal firing). The state of a neuronal unit is updated as a function of its current state and contributions from voltage-independent and voltage-dependent inputs. The voltage-independent input to unit i from unit j is:

$$A_{ij}^{VI}(t) = c_{ij}s_j(t) \quad (7)$$

where $s_j(t)$ is the activity of unit j , and c_{ij} is the connection strength from unit j to unit i . The voltage-independent post-synaptic influence, $POST_i^{VI}$, on unit i is calculated by summing over all the inputs onto unit i :

$$POST_i^{VI}(t) = \varphi(POST_i^{VI}(t-1)) + (1 - \varphi) \left(\sum_{l=1}^M \sum_{j=1}^{N_l} A_{ij}^{VI}(t) \right) \quad (8)$$

where M is the number of different anatomically defined connection types, N_l is the number of connections of type M projecting to unit i , and φ is the persistence of synaptic input.

The voltage-dependent input to unit i from unit j is:

$$A_{ij}^{VD}(t) = \Phi(POST_i^{VI}(t)) c_{ij} s_j(t), \text{ where } \Phi(x) \begin{cases} 0 & x < \sigma_i^{vdep} \\ x & \text{otherwise} \end{cases} \quad (9)$$

and σ_i^{vdep} is a threshold for the post-synaptic activity below which voltage-dependent connections have no effect.

The voltage-dependent post-synaptic influence on unit i , $POST_i^{VD}$, is given by:

$$POST_i^{VD}(t) = \varphi(POST_i^{VD}(t-1)) + (1 - \varphi) \left(\sum_{l=1}^M \sum_{j=1}^{N_l} A_{ij}^{VD}(t) \right) \quad (10)$$

The total post-synaptic influence on neuronal unit i is given by:

$$POST_i = \sum_{j=1}^{N_{VI}} POST_j^{VI}(t) + \sum_{k=1}^{N_{VD}} POST_k^{VD} \quad (11)$$

although it should be noted that Darwin VII possessed no voltage-dependent connections and therefore, for that model, that term was always zero.

The new activity of a neuronal unit i is determined by the following activation function:

$$s_i(t+1) = \phi \left(\tanh(g_i POST_i + \omega s_i(t)) \right), \text{ where } \phi(x) \begin{cases} 0 & x < \sigma_i^{fire} \\ x & \text{otherwise} \end{cases} \quad (12)$$

where g_i is a scaling constant and ω is a persistence constant for a given neuronal unit.

Synaptic strengths are subject to modification according to a synaptic rule that depends on the pre- and post-synaptic neuronal unit activities. Plastic synaptic connections are either value-independent ($EC_{IN} \rightarrow DG, CA3, CA1$; $DG \rightarrow CA3$; $CA3 \rightarrow CA1$; $CA1 \rightarrow EC_{OUT}$) or value-dependent ($CA1 \rightarrow S$, $CA1 \rightarrow M_{HDG}$). Both of these rules are based on a modified BCM learning rule (19), which has been shown to be equivalent to spike-timing dependent plasticity under certain conditions (20). Synapses between neuronal units with strongly correlated firing phases are potentiated and synapses between neuronal units with weakly correlated phases are depressed; the magnitude of change is determined as well by pre- and post-synaptic activities.

Value-independent synaptic changes in c_{ij} are given by:

$$\Delta c_{ij}(t+1) = \eta s_i(t) s_j(t) BCM(s_i) \quad (13)$$

where $s_i(t)$ and $s_j(t)$ are activities of post- and pre-synaptic units, respectively, and η is a fixed learning rate. The function BCM is implemented as a piecewise linear function, taking post-synaptic activity

as input, which is defined by a sliding threshold, θ , two inclinations (k_1, k_2) and a saturation parameter ρ ($\rho = 6$ throughout):

$$BCM(s) = \begin{cases} -k_1 s & s \leq \frac{\theta}{2} \\ k_1(s - \theta) & \frac{\theta}{2} < s \leq \theta \\ \frac{k_2}{\rho} \tanh \rho(s - \theta) & \text{otherwise} \end{cases} \quad (14)$$

The threshold is adjusted based on the post-synaptic activity:

$$\Delta\theta = 0.25(s^2 - \theta) \quad (15)$$

Value-independent plasticity was subject to weight normalization to prevent unbounded potentiation:

$$c_{ij} = \frac{c_{ij}}{\sqrt{\sum_{k=l}^K c_{kj}^2}} \quad (16)$$

where c_{ij} is a particular connection, and K is the total number of connections onto unit j .

The rule for value-dependent plasticity differs from the value-independent rule in that synaptic change is governed by the pre-synaptic activity, post-synaptic activity, and temporal difference derived from the value system. The synaptic change for Darwin VII value-dependent synaptic plasticity is given by:

$$\Delta c_{ij}(t+1) = V(t)(\eta s_i(t) BCM(s_i) - \epsilon(c_{ij}(t) - c_{ij}(0))) \quad (17)$$

and for Darwin X is given by:

$$\Delta c_{ij}(t+1) = \eta s_i(t) TD(t) BCM(s_i) - \epsilon(c_{ij}(t) - c_{ij}(0)) \quad (18)$$

where ϵ is a constant dictating how synaptic weights that are unpotentiated decay back towards their original value.

References

1. J. Krichmar, G. Edelman, *Cereb Cortex* **12**, 818 (2002).
2. J. L. Krichmar, D. A. Nitz, J. A. Gally, G. M. Edelman, *Proc Natl Acad Sci U S A* **102**, 2111 (2005).
3. J. L. Krichmar, A. K. Seth, D. A. Nitz, J. G. Fleischer, G. M. Edelman, *Neuroinformatics* **3**, 197 (2005).
4. W. Schultz, P. Dayan, P. R. Montague, *Science* **275**, 1593 (1997).
5. R. Morris, *J Neurosci Methods* **11**, 47 (1984).
6. R. U. Muller, J. B. J. Ranck, J. S. Taube, *Curr Opin Neurobiol* **6**, 196 (1996).
7. J. S. Taube, *Prog Neurobiol* **55**, 225 (1998).
8. A. Arleo, W. Gerstner, *Biol Cybern* **83**, 287 (2000).

9. J. P. Banquet, P. Gaussier, M. Quoy, A. Revel, Y. Burnod, *Neural Comput* **17**, 1339 (2005).
10. N. Burgess, J. G. Donnett, K. J. Jeffery, J. O'Keefe, *Philos Trans R Soc Lond B Biol Sci* **352**, 1535 (1997).
11. D. Amaral, N. Ishizuka, B. Claiborne, *Prog Brain Res* **83** (1990).
12. C. Bernard, H. Wheal, *Hippocampus* **4** (1994).
13. A. Treves, E. Rolls, *Hippocampus* **4**, 374 (1994).
14. T. F. Freund, G. Buzsaki, *Hippocampus* **6**, 347 (1996).
15. G. J. Mogenson, M. Nielsen, *Behav Neural Biol* **42**, 52 (1984).
16. A. M. Thierry, Y. Gioanni, E. Degenetais, J. Glowinski, *Hippocampus* **10**, 411 (2000).
17. A. G. Barto, R. S. Sutton, *Learning and Computational Neuroscience: Foundations of Adaptive Networks* (MIT Press, Cambridge, Massachusetts, 1990).
18. P. R. Montague, P. Dayan, T. J. Sejnowski, *J Neurosci* **16**, 1936 (1996).
19. E. L. Bienenstock, L. N. Cooper, P. W. Munro, *J Neurosci* **2**, 32 (1982).
20. E. M. Izhikevich, N. Desai, *Neural Comput* **15**, 1511 (1982).