Supporting Text

Neural Simulation

Darwin XI's 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 and a much less complex architecture. It 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 represents a local population of neurons. 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 of Darwin XI, 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 ≈200 ms of real time. The neural simulation was run on a Beowulf cluster containing 12 1.4 GHz 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 Darwin XI's position was recorded.

In the present experiments, the simulated nervous system contained 57 neural areas, 80,000 neuronal units, and ≈1.2 million synaptic connections. It included a visual system, a head direction system, a whisker system, a laser rangefinder system, a hippocampal formation, a basal forebrain, a value or reward system, and an action selection system. Fig. 2 shows a high-level diagram of the simulated nervous system including the various neural areas and the arrangement of synaptic connections. Specific parameters relating to each area and to patterns of connectivity are given in SI Tables 2 and 3.

Sensory Input

Visual images from Darwin XI's CCD camera were filtered for color and edges. The filtered output directly affected neural activity in area *V1*, which is composed of functionally segregated subareas for color and shape. The CCD camera sends 320 × 240 pixel RGB video images, via an RF transmitter, to a frame grabber attached to one of the workstations running the neural simulation. The image was spatially averaged to produce an 80 × 60 pixel image. Different sized Gabor filters (2 × 2, 4 × 14, 16 × 16, and 32 × 32) were used to detect vertical edges of varying widths. The output of the Gabor function mapped directly onto the neuronal units of the corresponding *V1* subarea (*V1-width2, V1-width4, V1-width16,* and *V1-width32*). The RGB video was transformed into YUV color space. Color filters in UV space (red, green, yellow, and blue) were applied to the image. The outputs of the color filters were mapped directly onto the neuronal units of *V1-red*, *V1-green, V1-blue,* and *V1-yellow*. *V1* color neuronal units projected nontopologically to inferotemporal cortex *IT*, and *V1* edge units projected retinotopically to parietal cortex *Pr* (see Fig. 2 and SI Table 3).

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 (1, 2). Neurons in these areas are often called head direction cells. Odometer information obtained from Darwin XI'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:

$$\left(\cos(HD_i - curr_heading) \right)^5 ; \qquad (1)$$

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 (see $HD \rightarrow ATN$ in SI Table 3 and Fig. 2) and to a motor area (see $HD \rightarrow M_{HDG}$ in SI Table 3 and Fig. 2) used for selecting a new heading (see below).

The whisker system consists of artificial whiskers and a model of somatosensory whisker barrel cortex. The whiskers produce activity in a set of thalamic lag units. Each lag cell is characterized by an internal state (s_i^{in}), an output (s_i), and a cell-specific lag parameter set to be $\psi_i = \frac{0.2}{i}, i \in \{1,2,\cdots,20\}$ for cell i in each whisker barrel. When triggered by a whisker deflection, the internal s_i^{in} state of cell i in the corresponding barrel increases at rate determined by ψ_i. When this internal state reaches a threshold, the cell begins to emit an output signal and s_i^{in} is reset to zero. Because of differences in ψ_i among lag cells, each whisker deflection evokes a wave of activity in the corresponding barrel, with some cells firing shortly after deflection and the remainder firing with gradually increasing delays.

Specifically, the internal state of each lag cell *i*, in the barrel corresponding to whisker *k*, is updated according to:

$$s_{ki}^{in}(t+1) = \begin{cases} 0.2; & s_{ki}^{in}(t) < 0.2,\ \overline{diff_k}(t) > 3.0 \\ 0; & s_{ki}^{in}(t) \geq \sigma_i^{fire} \\ (1+\psi_i)(s_{ki}^{in}(t)); & otherwise \end{cases} \tag{2}$$

where $\overline{diff_k}(t)$ is the difference between successive whisker readings averaged over the last four samples (a value exceeding 3.0 signifies a whisker deflection), and σ_i^{fire} is a firing threshold set to 0.3.

The output s_{ki} is calculated by using:

$$s_{ki}(t+1)=\begin{cases}\tanh\left(10(\omega_i(s_{ki}(t)))\right); & s_{ki}^{in}(t)<\sigma_i^{fire}\\ \tanh\left(10(\omega_i(s_{ki}(t))+(1-\omega_i)s_{ki}^{in}(t))\right); & otherwise\end{cases} \quad (3)$$

where $\omega_i = 0.8$ determines the persistence of unit activity. This value is fed as input into neuronal units in the corresponding barrels of *S1*.

Darwin XI's has three sensory whiskers on each side are arranged in a vertical stack. Whiskers are therefore referred to by the symbols LT, LM, LB to describe respectively the left top, left middle, and left bottom whiskers. Whisker lag and primary somatosensory areas in the simulation use these position suffixes to denote the whisker they receive input from (e.g., *WLLT, S1LT*).

A laser rangefinder (SICK LMS-200) produced distance-to-wall data, which was used by the CARMEN software package (http://carmen.sourceforge.net) to produce an estimate of device location. This location estimate was turned into activity in a topologically organized pseudocortical area (*SMAP*; see Fig. 2) that projected to EC_{in}. Each neuronal unit in *SMAP* responded preferentially to a particular place in the environment, and had a Gaussian tuning curve with 0.8 m standard deviation. Note that the laser data are being used merely as another sensory modality. Although the laser system produces an explicit estimate of location, this information is not necessary for producing place activity in the hippocampus.

Hippocampus

The architecture of the simulated hippocampal formation was based on rodent neuroanatomy. The input streams into the hippocampus are from the associative cortical areas in the simulation (see $ATN\rightarrow EC_{IN}$, $IT\rightarrow EC_{IN}$, $PR\rightarrow EC_{IN}$ in SI Table 3 and Fig. 2). 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} (see SI Tables 2 and 3). 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 (3-5). The perforant path projects mainly from entorhinal cortex to the dentate gyrus but also to the CA3 and CA1 subfields (see $EC_{IN\rightarrow}$ DG $EC_{IN\rightarrow}$ $CA3$, $EC_{IN\rightarrow}$ $CA3$ in SI Table 3 and Fig. 2). The mossy fibers (see $DG\rightarrow CA3$ in SI Table 3 and Fig. 2), Schaffer collaterals (see $CA3\rightarrow CA1$ in SI Table 3 and Fig. 2), and divergent projections from the hippocampus back to cortex (see $CA1\rightarrow EC_{OUT\rightarrow}$ ATN,IT,PR in SI Table 3 and Fig. 2) were also reflected in the neural simulation. Moreover, the prevalent recurrent connectivity found in the hippocampal formation was included in the model (see $EC_{IN\beta}\rightarrow EC_{OUT}$, $DG\rightarrow DG$, and $CA3\rightarrow CA3$ in SI Table 3 and Fig. 2).

There are distinct patterns of intrinsic and extrinsic, feedback and feedforward inhibitory connections in the hippocampal circuitry (5, 6). Feedback inhibitory connections (see $EC\rightarrow EC_{FB\rightarrow} EC$, $DG\rightarrow DG_{FB\rightarrow} DG$, $CA3\rightarrow CA3_{FB\rightarrow} CA3$, $CA1\rightarrow CA1_{FB\rightarrow} CA1$ in SI Table 3 and Fig. 2) and feedforward inhibitory connections (see $EC\rightarrow DG_{FF\rightarrow} DG$, $DG\rightarrow$ $CA3_{FF\rightarrow} CA3$, $CA3\rightarrow CA1_{FF\rightarrow} CA1$ in SI Table 3 and Fig. 2) 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:

$$BF(t) = theta(t \bmod 13); \qquad (4)$$

where *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 (see $BF\rightarrow$ $EC_{IN}, EC_{OUT}, DG, CA3, CA1$ in Fig. 2 and SI Table 3). The level of inhibition, which was adaptive, kept the activity in hippocampal regions within specific ranges:

$$\Delta sf_r(t) = (s_r(t) - tgt_r)$$
$$BF_r(t) = BF(t) + sf_r(t) \quad ; \qquad (5)$$

where r denotes the region (i.e., EC_{IN}, EC_{OUT}, DG, $CA3$, $CA1$), sf_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, tgt_r is the desired percentage of active units in area r (EC_{IN} = 10%, EC_{OUT} = 10%, DG = 20%, $CA3$ = 5%, and $CA1$ = 10%), and $BF_r(t)$ is the presynaptic neuronal unit activity for a connection from BF to hippocampus region r.

Action Selection

Activity in the simulated value system (Area S, Fig. 2) signals the occurrence of salient sensory events and this activity contributes to the modulation of value-dependent connection strengths in synaptic pathways ($CA1 \rightarrow S$ and $CA1 \rightarrow M_{HDG}$). The projection from our simulated $CA1$ to the value and goal decision areas is consistent with the connectivity between CA1 and nucleus accumbens and frontal areas (7, 8). Initially, S is activated by the hidden platform IR detector (see $T^{+} \rightarrow S$ in SI Table 3 and Fig. 2), causing potentiation of value-dependent connections. After experience, the value system could be activated by $CA1$. The magnitude of potentiation or depression is based on a neural implementation of a temporal difference learning rule (9, 10).

$$TD(t) = \begin{cases} T^{+}(t) - \overline{S(t-\tau)}; & T^{+} > 0 \\ \overline{S(t)} - \overline{S(t-\tau)}; & otherwise \end{cases} ; \qquad (6)$$

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 *Neuronal Dynamics* (below).

Darwin XI selected a new heading when reaching the choice point of the plus maze. The device stopped moving forward, and the camera was panned 90° to the left, and wait for three seconds, then pan the camera 90° to the right and wait for a further three secondsThe average activity of M_{HDG} was calculated during the wait periods. A softmax algorithm was used to create a probability distribution for choosing a new heading:

$$p(newhdg) = \frac{\exp\left(40\,\overline{M_{HDG}(newhdg)}\right)}{\sum_{h=hdg-60,hdg+60} \exp\left(40\,\overline{M_{HDG}(h)}\right)}; \tag{7}$$

where *newhdg* is a possible new heading for Darwin XI, $\overline{M_{HDG}(newhdg)}$ is the average activity of M_{HDG} at a possible new heading, *hdg* is the current heading, and *h* has two positions (current heading less 90° and current plus 90°).

Neuronal Dynamics and Synaptic Plasticity

A neuronal unit in Darwin XI 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 ≈200 ms. Synaptic connections between neural units, both within and between neuronal areas, are set to be either voltage-independent or -dependent, and either plastic or nonplastic (see Fig. 2 and SI Table 3). Voltage-independent connections provide synaptic input regardless of postsynaptic state. Voltage-dependent connections represent the contribution of receptor types (e.g., NMDA receptors) that require postsynaptic depolarization to be activated (11, 12).

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 (see Fig. 2). The voltage-independent input to unit i from unit j is:

$$A_{ij}^{VI}(t) = c_{ij}\, s_j(t), \tag{8}$$

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 postsynaptic influence, $POST_i^{VI}$, on unit i is calculated by summing over all of the inputs onto unit i:

$$POST_i^{VI}(t) = \varphi\left(POST_i^{VI}(t-1)\right) + (1-\varphi)\left(\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(A_{ij}^{VI}(t)\right)\right); \tag{9}$$

where M is the number of different anatomically defined connection types (see SI Table 3), 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:

$$A_{ij}^{VD}(t) = \Phi\left(POST_i^{VI}(t)\right) c_{ij}\, s_j(t), \text{ where } \Phi(x) = \begin{cases} 0; & x < \sigma_i^{vdep} \\ x; & otherwise \end{cases} \tag{10}$$

where σ_i^{vdep} is a threshold for the postsynaptic activity below which voltage-dependent connections have no effect (see SI Table 2).

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

$$POST_i^{VD}(t) = \varphi\left(POST_i^{VD}(t-1)\right) + (1-\varphi)\left(\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(A_{ij}^{VD}(t)\right)\right) \tag{11}$$

The total postsynaptic influence on neuronal unit i is given by:

$$POST_i = \sum_{j=1}^{N_{VI}} POST_j^{VI} + \sum_{k=1}^{N_{VD}} POST_k^{VD}; \tag{12}$$

The new activity is determined by the following activation function:

$$s_i(t+1) = \phi(\tanh(g_i(POST_i + \omega s_i(t)))), \text{ where } \phi(x) = \begin{cases} 0; & x < \sigma_i^{fire} \\ x; & otherwise \end{cases}; \tag{13}$$

Synaptic strengths are subject to modification according to a synaptic rule that depends on the preand postsynaptic neuronal unit activities. Plastic synaptic connections are either value-independent (see $EC_{IN\rightarrow} DG,CA3,CA1$; $DG \rightarrow CA3$; $CA3 \rightarrow CA1$; $CA1 \rightarrow EC_{OUT}$ in Fig. 2 and SI Table 3) or value-dependent (see $CA1 \rightarrow S$, $CA1 \rightarrow M_{HDG}$ in Fig. 2 and SI Table 3). Both of these rules are based on a modified BCM learning rule (13). 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 preand postsynaptic activities. The specific parameter settings for fine-scale synaptic connections are given in the equations below and SI Table 3.

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

$$\Delta c_{ij}(t+1) = \eta\, s_i(t) s_j(t) BCM(s_i), \tag{14}$$

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

$$BCM(s) = \begin{cases} -k_1 s; & s \le \theta/2 \\ k_1(s-\theta); & \theta/2 < s \le \theta \\ k_2 \tanh(\rho(s-\theta))/\rho; & otherwise \end{cases} \quad (15)$$

The threshold is adjusted based on the postsynaptic activity:

$$\Delta\theta = 0.25(s^2 - \theta) \quad (16)$$

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

$$c_{ij} = \frac{c_{ij}}{sqrt\left(\sum_{k=l}^{K} c_{kj}^2\right)}; \quad (17)$$

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 presynaptic activity, postsynaptic activity, and temporal difference derived from the value system. The synaptic change for value-dependent synaptic plasticity is given by:

$$\Delta c_{ij}(t+1) = \eta\, s_i(t) s_j(t) TD(t) - 0.002\left(c_{ij}(t) - c_{ij}(0)\right); \quad (18)$$

where $TD(t)$ is the temporal difference value at time t (see Eq. **6**). The second term in Eq. **18** generates synaptic decay toward the initial weight. This was necessary to enable the device to reverse its behavioral choice in the plus-maze task.

1. Muller RU, Ranck JB, Jr, Taube JS (1996) *Curr Opin Neurobiol* 6:196-206.

2. Taube JS (1998) *Prog Neurobiol* 55:225-256.

3. Treves A, Rolls ET (1994) *Hippocampus* 4:374-391.

4. Amaral DG, Ishizuka N, Claiborne B (1990) in *Progress in Brain Research*, eds Storm-Mathisen J, Zimmer J, Ottersen OP (Elsevier Science, Amsterdam), Vol 83, pp 1-11.

5. Bernard C, Wheal HV (1994) *Hippocampus* 4:497-529.

6. Freund TF, Buzsaki G (1996) *Hippocampus* 6:347-470.

7. Mogenson G, Nielsen M (1984) *Behav Neural Biol* 42:52-60.

8. Thierry A M, Gioanni Y, Degenetais E, Glowinski J (2000) *Hippocampus* 10:411-419.

9. Sutton RS, Barto AG (1990) in *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, ed Moore J (MIT Press, Cambridge, MA), pp 497-537.

10. Montague PR, Dayan P, Sejnowski TJ (1996) *J Neurosci* 16:1936-1947.

11. Grossberg S (1999) *Conscious Cognit* 8:1-44.

12. Wray J, Edelman GM (1996) *Cereb Cortex* 6:701-716.

13. Bienenstock EL, Cooper LN, Munro PW (1982) *J Neurosci* 2:32-48.