

A Dynamic Neural Network Model of Sensorimotor Transformations in the Leech

Shawn R. Lockery

Yan Fang

Terrence J. Sejnowski

Computational Neurobiological Laboratory,

Salk Institute for Biological Studies, Box 85800, San Diego, CA 92138 USA

Interneurons in leech ganglia receive multiple sensory inputs and make synaptic contacts with many motor neurons. These "hidden" units coordinate several different behaviors. We used physiological and anatomical constraints to construct a model of the local bending reflex. Dynamic networks were trained on experimentally derived input-output patterns using recurrent backpropagation. Units in the model were modified to include electrical synapses and multiple synaptic time constants. The properties of the hidden units that emerged in the simulations matched those in the leech. The model and data support distributed rather than localist representations in the local bending reflex. These results also explain counterintuitive aspects of the local bending circuitry.

1 Introduction

Neural network modeling techniques have recently been used to predict and analyze the connectivity of biological neural circuits (Zipser and Andersen 1988; Lehky and Sejnowski 1988; Anastasio and Robinson 1989). Neurons are represented as simplified processing units and arranged into model networks that are then trained to reproduce the input-output function of the reflex or brain region of interest. After training, the receptive and projective field of hidden units in the network often bear striking similarities to actual neurons and can suggest functional roles of neurons with inputs and outputs that are hard to grasp intuitively. We applied this approach to the local bending reflex of the leech, a three-layered, feedforward network comprising a small number of identifiable neurons whose connectivity and input-output function have been determined physiologically. We found that model local bending networks trained using recurrent backpropagation (Pineda 1987; Pearlmutter 1989) to reproduce a physiologically determined input-output function contained hidden units whose connectivity and temporal response properties closely resembled those of identified neurons in the biological network.

The similarity between model and actual neurons suggested that local bending is produced by distributed representations of sensory and motor information.

2 The Local Bending Reflex

In response to a mechanical stimulus, the leech withdraws from the site of contact (Fig. 1a). This is accomplished by contraction of longitudinal muscles beneath the stimulus and relaxation of longitudinal muscles on the opposite side of the body, resulting in a U-shaped local bend (Kristan 1982). The form of the response is independent of the site of stimulation: dorsal, ventral, and lateral stimuli produce an appropriately oriented withdrawal. Major input to the local bending reflex is provided by four pressure sensitive mechanoreceptors called P cells, each with a receptive field confined to a single quadrant of the body wall (Fig. 1b). Output to the muscles is provided by eight types of longitudinal muscle motor neurons, one to four excitatory and inhibitory motor neurons for each body wall quadrant (Stuart 1970; Ort et al. 1974). Motor neurons are connected by chemical and electrical synapses that introduce the possibility of feedback among the motor neurons.

Dorsal, ventral, and lateral stimuli each produces a pattern of P cell activation that results in a unique pattern of activation and inhibition of the motor neurons (Lockery and Kristan 1990a). Connections between sensory and motor neurons are mediated by a layer of interneurons (Kristan 1982). Nine types of local bending interneurons have been identified (Lockery and Kristan 1990b). These comprise the subset of the local bending interneurons that contributes to dorsal local bending because they are excited by the dorsal P cell and in turn excite the dorsal excitatory motor neuron. There appear to be no functional connections between interneurons. Other interneurons remain to be identified, such as those that *inhibit* the dorsal excitatory motor neurons.

Interneuron input connections were determined by recording the amplitude of the postsynaptic potential in an interneuron while each of the P cells was stimulated with a standard train of impulses (Lockery and Kristan 1990b). Output connections were determined by recording the amplitude of the postsynaptic potential in each motor neuron when an interneuron was stimulated with a standard current pulse. Most interneurons received substantial input from three or four P cells, indicating that the local bending network forms a distributed representation of sensory input (Fig. 1c).

3 Neural Network Model

Because sensory input is represented in a distributed fashion, most interneurons are active in all forms of local bending. Thus, in addition

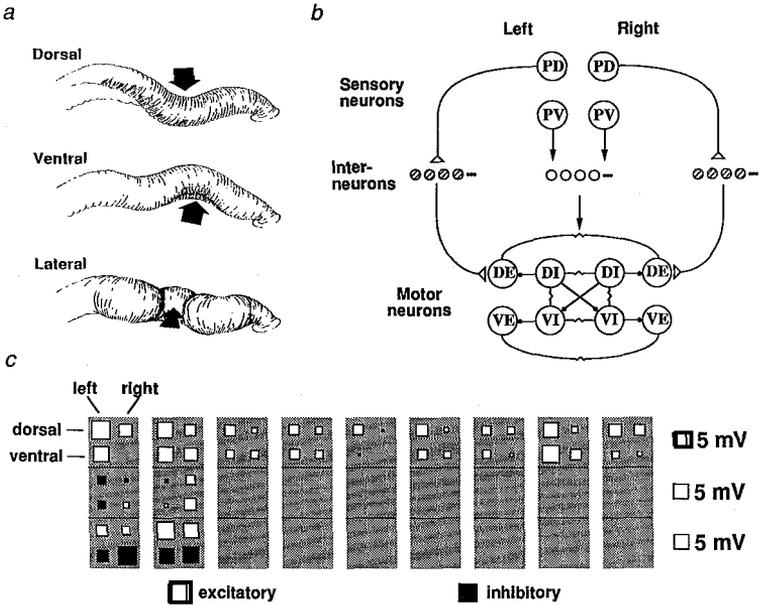


Figure 1: (a) Local bending behavior. Partial view of a leech in response to dorsal, ventral, and lateral stimuli. (b) Local bending circuit. The main input to the reflex is provided by the dorsal and ventral P cells (PD and PV). Control of local bending is largely provided by motor neurons whose field of innervation is restricted to single left-right, dorsal-ventral quadrants of the body; dorsal and ventral quadrants are innervated by both excitatory (DE and VE) and inhibitory (DI and VI) motor neurons. Motor neurons are connected by electrical synapses (resistor symbol) and excitatory (triangle) and inhibitory (filled circle) chemical synapses. Sensory input to motor neurons is mediated by a layer of interneurons. Interneurons that were excited by PD and that in turn excite DE have been identified (hatched); other types of interneurons remain to be identified (open). (c) Input and output connections of the nine types of dorsal local bending interneurons. Within each gray box, the upper panel shows input connections from sensory neurons, the middle panel shows output connections to inhibitory motor neurons, and the lower panel shows output connections to excitatory motor neurons. Box area is proportional to the amplitude of the connection determined from intracellular recordings of interneurons or motor neurons. White boxes indicate excitatory connections and black boxes indicate inhibitory connections. Blank spaces denote connections whose strength has not been determined.

to contributing to *dorsal* local bending, most interneurons are also active during *ventral* and *lateral* bending when some or all of their output effects are inappropriate to the observed behavioral response. This suggests that the inappropriate effects of the dorsal bending interneurons must be offset by other as yet unidentified interneurons and raises the possibility that local bending is the result of simultaneous activation of a population of interneurons with multiple sensory inputs and both appropriate and inappropriate effects on many motor neurons. It was not obvious, however, that such a population was sufficient, given the constraints imposed by the input-output function and connections known to exist in the network. The possibility remained that interneurons specific for each form of the behavior were required to produce each output pattern. To address this issue, we used recurrent backpropagation (Pearlmutter 1989) to train a dynamic network of model neurons (Fig. 2a). The network had four input units representing the four P cells, and eight output units representing the eight motor neuron types. Between input and output units was a single layer of 10 hidden units representing the interneurons. Neurons were represented as single electrical compartments with an input resistance and time constant. The membrane potential (V_i) of each neuron was given by

$$T_i dV_i/dt = -V_i + R_i(I_e + I_c)$$

where T_i and R_i are the time constant and input resistance of the neuron and I_e and I_c are the sum of the electrical and chemical synaptic currents from presynaptic neurons. Current due to electrical synapses was given by

$$I_e = \sum_j g_{ij} (V_j - V_i)$$

where g_{ij} is the coupling conductance between neuron i and j . To implement the delay associated with chemical synapses, synapse units (s-units) were inserted between pairs of neurons connected by chemical synapses. The activation of each s-unit was given by

$$T_{ij} dS_{ij}/dt = -S_{ij} + f(V_j)$$

where T_{ij} is the synaptic time constant and $f(V_j)$ was a physiologically determined sigmoidal function ($0 \leq f \leq 1$) relating pre- and postsynaptic membrane potential at an identified monosynaptic connection in the leech (Granzow et al. 1985). Current due to chemical synapses was given by

$$I_c = \sum_j w_{ij} S_{ij}$$

where w_{ij} is the strength of the chemical synapse between units i and j . Thus synaptic current is a graded function of presynaptic voltage, a common feature of neurons in the leech (Friesen 1985; Granzow et al. 1985;

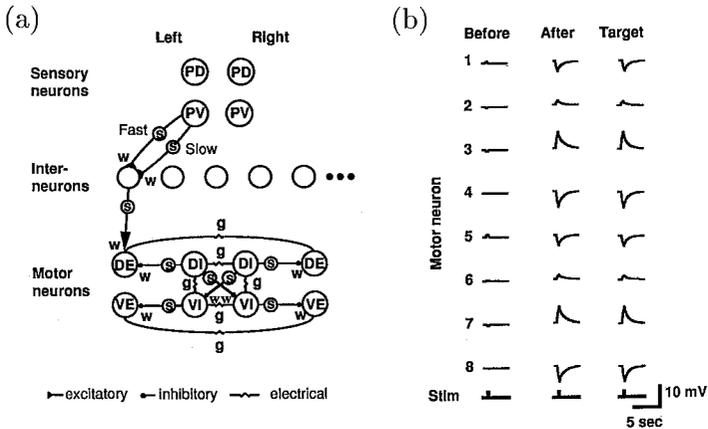


Figure 2: (a) The local bending network model. Four sensory neurons were connected to eight motor neurons via a layer of 10 interneurons. Neurons were represented as single electrical compartments whose voltage varied as a function of time (see text). Known electrical and chemical connections among motor neurons were assigned fixed connection strengths (g 's and w 's in the motor layer) determined from intracellular recordings. Interneuron input and output connections were adjusted by recurrent backpropagation. Chemical synaptic delays were implemented by inserting s -units between chemically connected pairs of neurons. S -units with different time constants were inserted between sensory and interneurons to account for fast and slow components of synaptic potentials recorded in interneurons (see Fig. 4). (b) Output of the model network in response to simultaneous activation of both PDs (stim). The response of each motor neuron (rows) is shown before and after training. The desired responses from the training set are shown on the right for comparison (target).

Thompson and Stent 1976) and other invertebrates (Katz and Miledi 1967; Burrows and Siegler 1978; Nagayama and Hisada 1987).

Chemical and electrical synaptic strengths between motor neurons were determined by recording from pairs of motor neurons and were not adjusted by the training algorithm. Interneuron input and output connections were given small initial values that were randomly assigned and subsequently adjusted during training. During training, input connections were constrained to be positive to reflect the fact that only excitatory interneuron input connections were seen (Fig. 1c), but no constraints were placed on the number of input or output connections. Synaptic time constants were assigned fixed values. These were adjusted by hand to fit

the time course of motor neuron synaptic potentials (Lockery and Kristan 1990a), or determined from pairwise motor neuron recordings (Granzow et al. 1985).

4 Results

Model networks were trained to produce the amplitude and time course of synaptic potentials recorded in all eight motor neurons in response to trains of P cell impulses (Lockery and Kristan 1990a). The training set included the response of all eight motor neurons when each P cell was stimulated alone and when P cells were stimulated in pairs. After 6,000–10,000 training epochs, the output of the model closely matched the desired output for all patterns in the training set (Fig. 2b). To compare interneurons in the model network to actual interneurons, simulated physiological experiments were performed. Interneuron input connections were determined by recording the amplitude of the postsynaptic potential in a model interneuron while each of the P cells was stimulated with a standard current pulse. Output connections were determined by recording the amplitude of the postsynaptic potential in each motor neuron when an interneuron was stimulated with a standard current pulse. Model and actual interneurons were compared by counting the number of input and output connections that were sufficiently strong to produce a postsynaptic potential of 0.5 mV or more in response to a standard stimulus. Model interneurons (Fig. 3a), like those in the real network (Fig. 1c), received three or four substantial connections from P cells and had significant effects on most of the motor neurons (Fig. 3b and c). Most model interneurons were therefore active during each form of the behavior and the output connections of the interneurons were only partially consistent with each form of the local bending response. Thus the appropriate motor neuron responses were produced by the summation of many appropriate and inappropriate interneuron effects. This result demonstrates that the apparently appropriate and inappropriate effects of dorsal local bending interneurons reflect the contribution these interneurons make to other forms of the local bending reflex.

There was also agreement between the time course of the response of model and actual interneurons to P cell stimulation (Fig. 4). In the actual network, interneuron synaptic potentials in response to trains of P cell impulses had a fast and slow component. Some interneurons showed only the fast component, some only the slow, and some showed both components (mixed). Although no constraints were placed on the temporal response properties of interneurons, the same three types of interneuron were found in the model network. The three different types of interneuron temporal response were due to different relative connection strengths of fast and slow s-units impinging on a given interneuron (Fig. 2a).

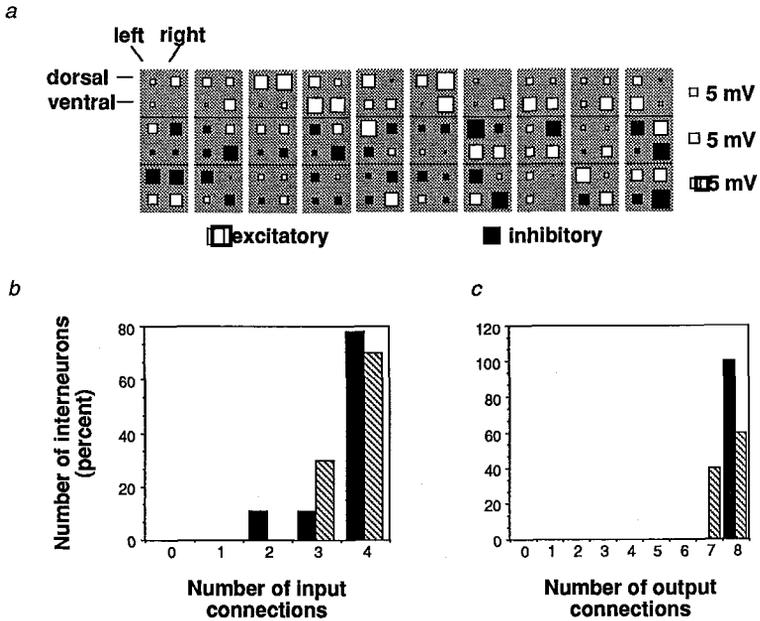


Figure 3: (a) Input and output connections of model local bending interneurons. Model interneurons, like the actual interneurons, received substantial inputs from three or four sensory neurons and had significant effects on most of the motor neurons. Symbols as in Figure 1c. (b,c) Number of interneurons in model (hatched) and actual (solid) networks with the indicated number of significant input and output connections. Input connections (b) and output connections (c) were considered significant if the synaptic potential in the postsynaptic neuron was greater than 0.5 mV. Counts are given in percent of all interneurons because model and actual networks had different numbers of interneurons.

5 Discussion

Our results show that the network modeling approach can be adapted to models with more realistic neurons and synaptic connections, including electrical connections, which occur in both invertebrates and vertebrates. The qualitative similarity between model and actual interneurons demonstrates that a population of interneurons resembling the identified dorsal local bending interneurons could mediate local bending in a distributed

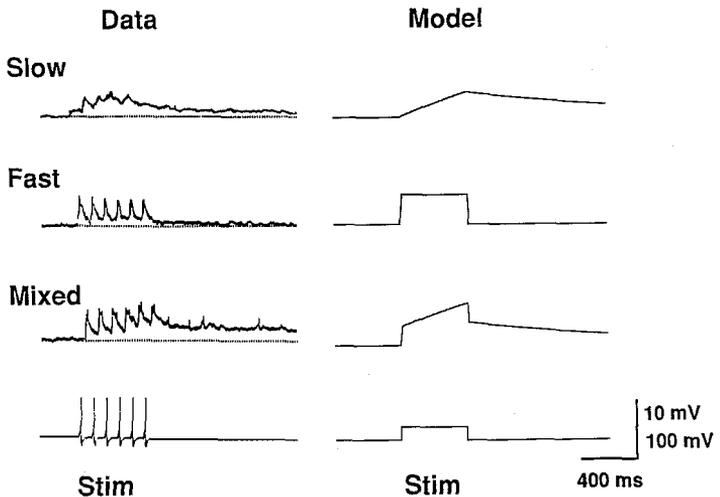


Figure 4: Actual (data) and simulated (model) synaptic potentials recorded from three types of interneuron. Actual synaptic potentials were recorded in response to a train of P cell impulses (stim). Simulated synaptic potentials were recorded in response to a pulse of current in the P cell, which approximates a step change in P cell firing frequency.

processing system without additional interneurons specific for different forms of local bending. Interneurons in the model also displayed the diversity in temporal responses seen in interneurons in the leech. This represents an advance over our previous model in which temporal dynamics were not represented (Lockery et al. 1989). Clearly, the training algorithm did not produce exact matches between model and actual interneurons, but this was not surprising since the identified local bending interneurons represent only a subset of the interneurons in the reflex. More exact matches could be obtained by using two pools of model interneurons, one to represent identified neurons, the other to represent unidentified neurons. Model neurons in the latter pool would constitute testable physiological predictions of the connectivity of unidentified local bending interneurons.

Acknowledgments

Supported by the Bank of America-Giannini Foundation, the Drown Foundation, and the Mathers Foundation.

References

- Anastasio, T., and Robinson, D. A. 1989. Distributed parallel processing in the vestibulo-oculomotor system. *Neural Comp.* **1**, 230–241.
- Burrows, M., and Siegler, M. V. S. 1978. Graded synaptic transmission between local interneurons and motor neurons in the metathoracic ganglion of the locust. *J. Physiol.* **285**, 231–255.
- Friesen, W. O. 1985. Neuronal control of leech swimming movements: Interactions between cell 60 and previously described oscillator neurons. *J. Comp. Physiol.* **156**, 231–242.
- Granzow, B., Friesen, W. O., and Kristan, W. B., Jr. 1985. Physiological and morphological analysis of synaptic transmission between leech motor neurons. *J. Neurosci.* **5**, 2035–2050.
- Katz, B., and Miledi, R. 1967. Synaptic transmission in the absence of nerve impulses. *J. Physiol.* **192**, 407–436.
- Kristan, W. B., Jr. 1982. Sensory and motor neurons responsible for local bending in the leech. *J. Exp. Biol.* **96**, 161–180.
- Lehky, S. R., and Sejnowski, T. J. 1988. Network model of shape-from-shading: Neural function arises from both receptive and projective fields. *Nature (London)* **333**, 452–454.
- Lockery, S. R., and Kristan, W. B., Jr. 1990a. Distributed processing of sensory information in the leech. I. Input-output relations of the local bending reflex. *J. Neurosci.* **10**, 1811–1815.
- Lockery, S. R., and Kristan, W. B., Jr. 1990b. Distributed processing of sensory information in the leech. II. Identification of interneurons contributing to the local bending reflex. *J. Neurosci.* **10**, 1815–1829.
- Lockery, S. R., Wittenberg, G., Kristan, W. B., Jr., and Cottrell, W. G. 1989. Function of identified interneurons in the leech elucidated using neural networks trained by back-propagation. *Nature (London)* **340**, 468–471.
- Nagayama, T., and Hisada, M. 1987. Opposing parallel connections through crayfish local nonspiking interneurons. *J. Comp. Neurol.* **257**, 347–358.
- Ort, C. A., Kristan, W. B., Jr., and Stent, G. S. 1974. Neuronal control of swimming in the medicinal leech. II. Identification and connections of motor neurons. *J. Comp. Physiol.* **94**, 121–154.
- Pearlmutter, B. A. 1989. Learning state space trajectories in recurrent neural networks. *Neural Comp.* **1**, 263–269.
- Pineda, F. 1987. Generalization of backpropagation to recurrent neural networks. *Phys. Rev. Lett.* **19**, 2229–2232.
- Stuart, A. E. 1970. Physiological and morphological properties of motoneurons in the central nervous system of the leech. *J. Physiol.* **209**, 627–646.
- Thompson, W. J., and Stent, G. S. 1976. Neuronal control of heartbeat in the medicinal leech. *J. Comp. Physiol.* **111**, 309–333.
- Zipser, D., and Andersen, R. A. 1988. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature (London)* **331**, 679–684.