What are the Computations of the Cerebellum, the Basal Gangila, and the Cerebral Cortex?

Kenji Doya
doya@erato.atr.co.jp
Kawato Dynamic Brain Project, ERATO
Japan Science and Technology Corporation
2-2 Hikaridai, Seika, Soraku, Kyoto 619-0288, Japan
Phone: +81-774-95-1210. Fax: +81-774-95-3001.

April 26th, 1999
to appear in *Neural Networks*

Abstract

The classical notion that the cerebellum and the basal ganglia are dedicated to motor control is under dispute given increasing evidence of their involvement in non-motor functions. Is it then impossible to characterize the functions of the cerebellum, the basal ganglia and the cerebral cortex in a simplistic manner? This paper presents a novel view that their computational roles can be characterized not by asking what are the "goals" of their computation, such as motor or sensory, but by asking what are the "methods" of their computation, specifically, their learning algorithms. There is currently enough anatomical, physiological, and theoretical evidence to support the hypotheses that the cerebellum is a specialized organism for supervised learning, the basal ganglia are for reinforcement learning, and the cerebral cortex is for unsupervised learning.

This paper investigates how the learning modules specialized for these three kinds of learning can be assembled into goal-oriented behaving systems. In general, supervised learning modules in the cerebellum can be utilized as "internal models" of the environment. Reinforcement learning modules in the basal ganglia enable action selection by an "evaluation" of environmental states. Unsupervised learning modules in the cerebral cortex can provide statistically efficient representation of the states of the environment and the behaving system. Two basic action selection architectures are shown, namely, reactive action selection and predictive action

selection. They can be implemented within the anatomical constraint of the network linking these structures. Furthermore, the use of the cerebellar supervised learning modules for state estimation, behavioral simulation, and encapsulation of learned skill is considered. Finally, the usefulness of such theoretical frameworks in interpreting brain imaging data is demonstrated in the paradigm of procedural learning.

1 Introduction

Traditionally, the cerebellum and the basal ganglia were thought to be dedicated to the control of movement. However, this idea is rapidly losing ground to the corpus of evidence that suggests their involvement in non-motor tasks (Leiner et al., 1993; Middleton and Strick, 1994; Desmond and Fiez, 1998; Brown et al., 1997; Lawrence et al., 1998). Despite the demise of the motor control-only dogma, the activity of the cerebellum and the basal ganglia in brain imaging studies tends to be neglected or under-evaluated, often as the artifacts of motor components in the task response. This negligence is not only due to the wide prevalence of the old theory, but also due to the paucity of alternative theories that would enable us to comprehend the way the cerebellum and the basal ganglia participate in sensory or cognitive tasks (Ito, 1993; Houk, 1997).

One characteristic feature of the cerebellum is its highly uniform anatomical organization (Llinas and Walton, 1998). The characteristic multiple inhibitory pathways are also present throughout the parallel circuit of the basal ganglia (Wilson, 1998). The cerebral neocortex also has a characteristic six-layer organization (Douglas and Martin, 1998). These anatomical features, which are preserved throughout each of these structures, suggest that the cerebellum, the basal ganglia, and the cerebral cortex are each specialized for a certain kind of computation. However, the question about their unique "purposes" of computation, such as motor control, sensory acquisition, or cognitive processing, is ill-posed because all of these functions depend on each other under a normal behavioral context (Bower, 1997). Furthermore, the cerebellum and the basal ganglia are reciprocally connected to the cerebral cortex (Figure 1), and the connected areas tend to be simultaneously active, making it difficult to differentiate their roles by a mere observation of their activity.

The approach taken in this paper is to ask what are the unique "methods" of computation, namely, the learning algorithms, for which the cerebellum, the basal ganglia, and the cerebral cortex are specialized. The specific hypotheses considered in this paper are that the cerebellum is a specialized organism for *supervised learning*, the basal ganglia is for *reinforcement learning*, and the cerebral cortex is for *unsupervised learning*. In this framework, a particular function, such as the control of arm movement, can be realized by a global network combining different learning modules in the cerebellum, the basal ganglia, and the

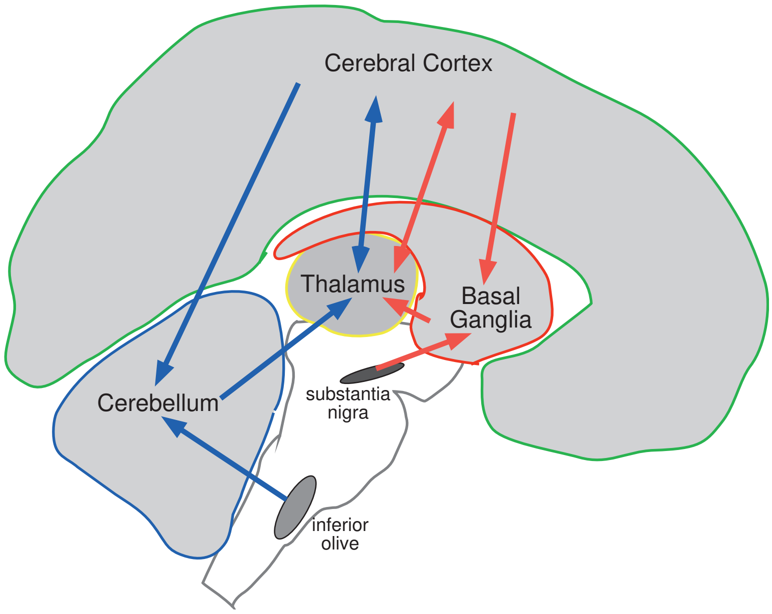

Figure 1: Global network linking the cerebellum, the basal ganglia, and the cerebral cortex.

cerebral cortex. Moreover, learning modules in the same structure, for example, different parts of the cerebellum, can be used for different behavioral purposes, for example, for motor control and cognitive processing, depending on where their outputs are directed (Ito, 1993; Middleton and Strick, 1996; Wolpert et al., 1998).

In Section 2, the goals and the algorithms of supervised, reinforcement, and unsupervised learning paradigms are summarized. Section 3 summarizes the anatomical, physiological, and theoretical evidence that suggests that the cerebellum, the basal ganglia, and the cerebral cortex are specialized, respectively, in the three learning paradigms reviewed in Section 3. In Section 4, a number of global architectures are considered in which supervised, reinforcement, and unsupervised learning modules are assembled for goal-directed behaviors. It is also investigated whether these architectures can be implemented within the anatomical and physiological constraints. In Section 5, experimental data on the brain activity during procedural learning tasks are considered in light of the behavioral architectures considered in Section 4. Finally, Section 6 provides discussions on the issues concerning the learning of global architectures and the regulation of global parameters.

2 Three Basic Learning Paradigms

This section presents a summary of the formulations and the algorithms of the three major learning paradigms considered in the computational learning theories, namely, supervised

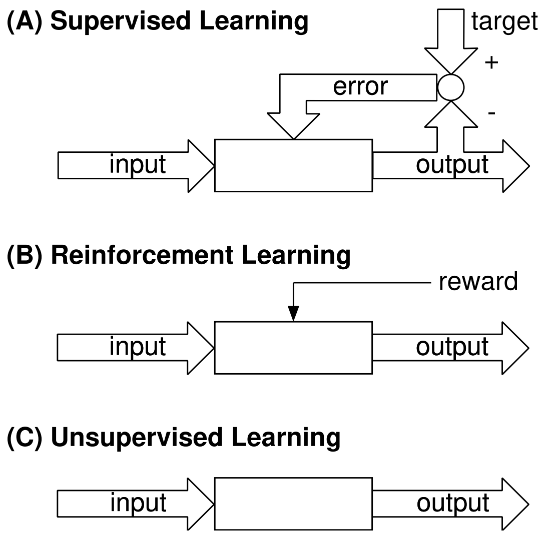

Figure 2: Three basic learning paradigms. (A) Supervised learning by error vectors. (B) Reinforcement learning by scalar reward signal. (C) Unsupervised learning by statistics of input signal itself.

learning, reinforcement learning, and unsupervised learning (Figure 2). This classification is due to the nature of the *teaching signals* that guide learning: directional *error vectors* in supervised learning, scalar *rewards*, or *reinforcement* signals, in reinforcement learning, and *none* in unsupervised learning.

2.1 Supervised Learning

In the supervised learning paradigm (Figure 2A), the goal is to construct an input-output mapping

$$\mathbf{y} = F(\mathbf{x}) \tag{1}$$

that predicts the output $\mathbf{y} = (y_1, ..., y_m)'$ for an input data point $\mathbf{x} = (x_1, ..., x_n)'$. The mapping is found from the examples of the *desired output* $(\hat{\mathbf{y}}(1), \hat{\mathbf{y}}(2), ...)$ at the input data points $(\mathbf{x}(1), \mathbf{x}(2), ...)$ in order to minimize the expected output error, such as that in a Euclid norm

$$E_{\mathbf{x}}[||\hat{\mathbf{y}} - \mathbf{y}||^2]. \tag{2}$$

Since the true distribution of the input $p(\mathbf{x})$ is usually unknown, it is approximated by minimizing the sum of squared errors at sample data points

$$E = \sum_t ||\hat{\mathbf{y}}(t) - \mathbf{y}(t)||^2 = \sum_t ||\hat{\mathbf{y}}(t) - F(\mathbf{x}(t); \mathbf{w})||^2 \tag{3}$$

under a certain constraint on the mapping F, such as smoothness (see Bishop (1995) for more details).

For example, when the input-output mapping is represented as

$$y_i(t) = \sum_{j=1}^{n} w_{ij} x_j(t), \tag{4}$$

where w_{ij} is a connection weight, the supervised learning algorithm is given by the gradient descent of the sample error

$$\Delta w_{ij} \propto -\frac{\partial E}{\partial w_{ij}} = (\hat{y}_i(t) - y_i(t))\ x_j(t), \tag{5}$$

where $\hat{y}_i(t) - y_i(t)$ is the *output error.*

In summary, supervised learning is characterized by the parameter update based on the correlation between the output error and the presynaptic input.

2.2 Reinforcement Learning

In the reinforcement learning paradigm (Figure 2B), a *learning agent* takes an *action* $\mathbf{u}(t) \in R^m$ in response to the *state* $\mathbf{x}(t) \in R^n$ of the environment, which results in the change of the state

$$x(t+1) = F(\mathbf{x}(t), \mathbf{u}(t)), \tag{6}$$

and the delivery of reinforcement signal, or a *reward*

$$r(t+1) = R(\mathbf{x}(t), \mathbf{u}(t)). \tag{7}$$

The goal is to find a *policy*

$$\mathbf{u}(t) = G(\mathbf{x}(t)), \tag{8}$$

which maximizes the cumulative sum of the rewards.

In general, current action $\mathbf{u}(t)$ affects all future states and accordingly all future rewards. The maximization is realized by the use of the *value function* of the states

$$V(\mathbf{x}) = E[r(t+1) + \gamma r(t+2) + \gamma^2 r(t+3) + ...], \tag{9}$$

where the right hand side is evaluated for the dynamics (6), (7), and (8) with the initial condition $\mathbf{x}(t) = \mathbf{x}$. The two basic ingredients in reinforcement learning are, first, the estimation of the value function, and then, the improvement of the policy using the value function (see Sutton and Barto (1998) for more information).

The basic algorithm for learning the value function is to minimize the *temporal difference (TD) error* of the reward prediction

$$\delta(t) = r(t) + \gamma V(\mathbf{x}(t)) - V(\mathbf{x}(t-1)), \tag{10}$$

which signals the inconsistency of the current estimate of the value function. For example, when the value function is represented as

$$V(t) = \sum_{j=1}^{n} v_j x_j(t), \tag{11}$$

the learning algorithm for the weight v_j is given by

$$\Delta v_j \propto \delta(t) x_j(t-1). \tag{12}$$

A simple way of improving the policy is to take a stochastic action

$$u_i(t) = g(\sum_{j=1}^{n} w_{ij} x_j(t) + \nu_i(t)), \tag{13}$$

where $g()$ is a gain function and $\nu_i(t)$ is a noise term. The TD error $\delta(t)$ as defined in (10) then signals the unexpected delivery of the reward $r(t)$ or the increase in the state value $V(\mathbf{x}(t))$ above expectation, possibly due to the previous choice of action $u_i(t-1)$. The learning algorithm for the action weight w_{ij} is given by

$$\Delta w_{ij} \propto \delta(t)(u_i(t-1) - \bar{u}_i) x_j(t-1), \tag{14}$$

where $\bar{u}_i$ is the average level of the action output.

Thus, the TD error $\delta(t)$, which signals the error in reward prediction, works as the main teaching signal in both learning of the value function and the selection of actions,

2.3 Unsupervised Learning

In the unsupervised learning paradigm (Figure 2C), only a set of input data $(\mathbf{x}(1), \mathbf{x}(2), ...) \in R^n$ is given and the goal is to construct a mapping so that the output $(\mathbf{y}(1), \mathbf{y}(2), ...) \in R^m$ fully characterizes the statistical properties of the input.

A typical formulation is a maximization of the mutual information between the input and the output

$$H(\mathbf{x}; \mathbf{y}) = H(\mathbf{x}) - H(\mathbf{x}|\mathbf{y}), \tag{15}$$

where H denotes the *entropy* $H(\mathbf{x}) = E[-\log p(\mathbf{x})]$. The mutual information quantifies the decrease in uncertainty about input $\mathbf{x}$ by knowing output $\mathbf{y}$ (see Becker (1995) for more details).

For example, a typical unsupervised learning algorithm is derived from an objective function

$$E = ||\mathbf{x}(t) - W'\mathbf{y}(t)||^2 + \sum_{i=1}^{m} |y_i(t)|, \tag{16}$$

where W is the input-output weight matrix and W' denotes its transpose. The first term represents the "input reconstruction error" that is minimized when the mutual information is maximized. The second term provides a sparseness constraint that encourages the majority of output units to stay close to zero (Harpur and Prager, 1995; Olshausen and Field, 1996).

Output $\mathbf{y}$ is determined as the fixed point solution of the relaxation dynamics

$$\dot{\mathbf{y}} \propto -\frac{\partial E}{\partial \mathbf{y}} = W\mathbf{x} - WW'\mathbf{y} - \mathrm{sign}(\mathbf{y}). \tag{17}$$

The weights are updated by the gradient descent

$$\Delta W \propto -\frac{\partial E}{\partial W} = \mathbf{y}(\mathbf{x} - W'\mathbf{y})' = \mathbf{y}\mathbf{x}' - \mathbf{y}\mathbf{y}'W, \tag{18}$$

which is a Hebbian potentiation and activity-dependent synaptic decay.

In summary, unsupervised learning is characterized by the relaxation dynamics for determining the output as well as the Hebbian synaptic rule under a certain regularization.

3 Learning Paradigms of the Cerebellum, the Basal Ganglia, and the Cerebral Cortex

This section summarizes the anatomical and physiological evidence supporting the hypothesis that the cerebellum, the basal ganglia, and the cerebral cortex are specialized, respectively, in supervised, reinforcement, and unsupervised learning.

3.1 Supervised Learning in the Cerebellum

Anatomy: The circuit of the cerebellum is characterized by the nearly feed-forward structure (Figure 3) with massive synaptic convergence of granule cell axons (parallel fibers) onto Purkinje cells. There are two major input pathways: the *mossy fiber* input and the *climbing fiber* input from the inferior olive (see Llinas and Walton (1998) and DeZeeuw et al. (1998) for details). The mossy fiber input, which carries both sensory afferent and cerebral efferent signals, is relayed by a massive number of granule cells, each of which combines different mossy fiber inputs. The parallel fiber input and the climbing fiber input converge at the Purkinje cells, where each Purkinje cell receives approximately 200,000 parallel fiber inputs and only one climbing fiber input. The output neurons of the cerebellum, which are located in the deep cerebellar nuclei, receive inhibitory input from Purkinje cells as well as the excitatory mossy fiber input.

The connection between the cerebellar cortex, the cerebellar nuclei, and the inferior olive is highly topographic. There is a modular structure called "microzones," each of which consists of approximately 3,000 Purkinje cells. The outputs of the cerebellar nuclei are

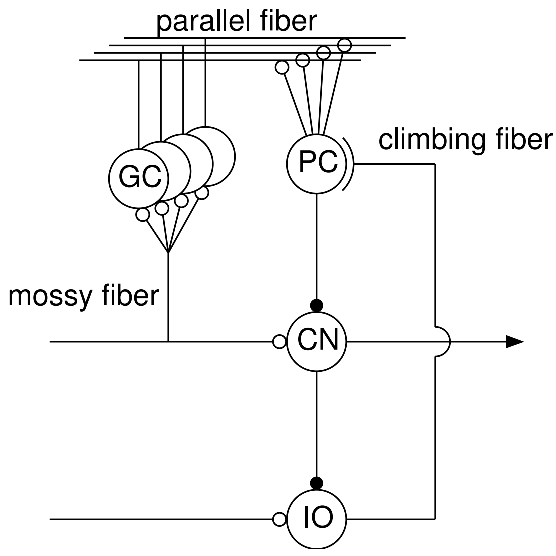

Figure 3: Diagram of cerebellar circuit. GC, granule cells; PC, Purkinje cells; CN, deep cerebellar nuclei; IO, inferior olive; $\circ$, excitatory connection; $\bullet$, inhibitory connection.

directed to the oculomotor system, the spinal motor pathway, and through the thalamus to certain areas in the cerebral cortex, depending on the locations within the cerebellar nuclei (Middleton and Strick, 1998).

Physiology: The responses of a Purkinje cell to parallel fiber inputs and climbing fiber input are discerned as two differently shaped spikes, which are called simple spikes and complex spikes. The simple spike response of the Purkinje cell, which is due to the parallel fiber input, encodes movement related signals, such as the acceleration and velocity components of eye movement command (Shidara et al., 1993; Gomi et al., 1998). The complex spike response, which is generated by a single spike of the climbing fiber, is characterized by its very low frequency, usually a few spikes per second. A statistical analysis has shown that the climbing fiber input best encodes the errors in movement, for example, the retinal slip signal in eye movement control (Kobayashi et al., 1998). The complex spikes are also related to movement errors in arm reaching movement, particularly toward the end of the movement (Kitazawa et al., 1998).

Coincident activation of the parallel fiber and climbing fiber inputs induces the *long-term depression (LTD)* of the parallel fiber synapse onto the Purkinje cell (Ito et al., 1982). The cellular and molecular mechanisms of the LTD have been studied in detail (Daniel et al., 1998).

Theoretical models: The results of these studies suggest that the circuit of the cerebellum is capable of implementing the supervised learning paradigm (Marr, 1969; Albus, 1971; Ito, 1984). The massive number of granule cells can work as expansion encoders of the mossy

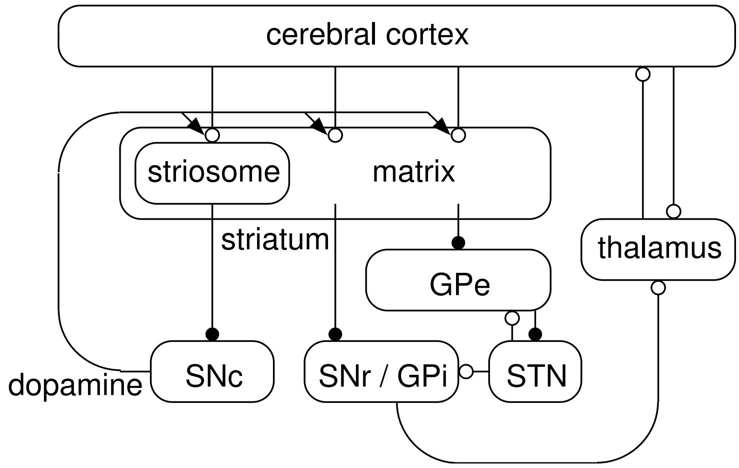

Figure 4: Diagram of neural circuit of the basal Ganglia. SNc, substantia nigra, pars compacta; SNr, substantia nigra, pars reticulata; GPi, internal segment of globus pallidus; GPe, external segment of globus pallidus; STN, subthalamic nucleus; Th, thalamus; $\circ$, excitatory connection; $\bullet$, inhibitory connection.

fiber input signal. Their outputs are linearly combined by a Purkinje cell, as in (4), and the synaptic weights are updated by (5) with the error signal $\hat{y}_i(t) - y_i(t)$ carried by the climbing fiber and the input $x_j(t)$ carried by the parallel fibers (Kawato and Gomi, 1992a; Wolpert et al., 1998).

The involvement of the cerebellum in error-driven learning behaviors, such as eye-movement control and conditioned eye-blinking, has been experimentally demonstrated (Ito, 1984; Thompson and Krupa, 1994), although there is still controversy about whether the LTD in the parallel fiber-Purkinje cell synapse is the single major substrate of learning (Raymond et al., 1996). Houk et al. (1996) have provided a review of models of the cerebellum at different levels of abstraction.

A number of simulation and robotic experiments have confirmed the usefulness of supervised learning in cerebellum-like networks in motor control tasks (Fujita, 1982; Kawato et al., 1987; Gomi and Kawato, 1992; Kettner et al., 1997; Schweighofer et al., 1998). An important issue in supervised learning is the time delay in the delivery of the error signal, which can be circumvented by introducing an “eligibility trace” for synaptic plasticity (Houk and Alford, 1996; Schweighofer et al., 1996; Kettner et al., 1997).

3.2 Reinforcement Learning in the Basal Ganglia

Anatomy: The circuit of the basal ganglia is characterized by multiple inhibitory pathways, as illustrated in Figure 4. The striatum (the caudate nucleus and the putamen) receives the main input from the cerebral cortex. The striatum consists of two compartments: the *striosome* that sends output to the dopamine neurons in SNc and the *matrix* that sends output to GPi and SNr (see the legend in Figure 4 for complete terms). There are both direct and indirect pathways that work in a complimentary manner. The dopamine neurons in SNc send projection back to both compartments of the striatum. Those dopaminergic terminal are found in the presynaptic terminals of the cortico-striatal connection. The outputs of GPi and SNr are directed through the thalamus to specific areas in the cerebral cortex. There are multiple loops starting from the cerebral cortex, through the basal ganglia and the thalamus, and back to the cerebral cortex (Alexander and Crutcher, 1990). It has recently been shown that the projection from the GPi through the thalamus to the cortex is highly topographic, and thus the basal ganglia are composed of multiple "output channels" (Middleton and Strick, 1994; Middleton and Strick, 1996).

Physiology: Data from neuronal recording and lesion studies indicate that the basal ganglia are involved in learning and execution of goal-directed, sequential behavior (Graybiel, 1995; Miyachi et al., 1997). Neurons in the striatum and the global pallidus show activities at different timings for the sensory-motor behaviors (Schultz et al., 1995). "Tonically active neurons" in the striatum show a phasic pause of spiking when an animal detects sensory cues that signals the delivery of a reward (Aosaki et al., 1994). The dopamine neurons in SNc, which send output to the striatum, show phasic increase in firing when an animal is given an unexpected reward or a sensory cue that signals the delivery of a reward in the near future (Schultz et al., 1993; Schultz, 1998).

The plasticity of the cortical input to the striatum is known to be modulated by the dopamine. It was shown that the coincident cortical input and the depolarization in a striatal neuron induce long-term depression (LTD) of the cortico-striatal synapse, provided that there are enough level of dopamine (Calabresi et al., 1996). Furthermore, long-term potentiation (LTP) instead of LTD is observed when the coincidence of cortical input and postsynaptic depolarization is also associated with a phasic dopamine release (Wickens et al., 1996).

Theoretical models: Based on the reward predicting activity of dopamine neurons and the change of behavior in the course of task learning, it has been suggested that the basal ganglia is involved in reinforcement learning, in which the dopamine neuron activity encoding the reward prediction error $\delta(t)$ in (10) (Montague et al., 1996; Schultz et al., 1997). More specifically, it is suggested that the striosome compartment works as the value prediction mechanism while the matrix compartment works as the action selection mechanism (Barto, 1995; Houk et al., 1995).

The capability of reinforcement learning models of the basal ganglia to replicate sequence learning behaviors has been demonstrated (Dominey et al., 1995; Nakahara et al., 1997; Nakahara et al., 1998; Suri and Schultz, 1998; Berns and Sejnowski, 1998).

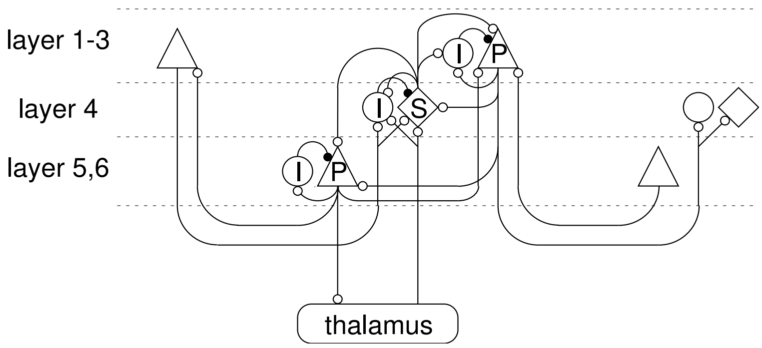

Figure 5: Diagrams of neural circuit of the cerebral cortex. P, pyramidal neurons; S, spiny stellate neurons; I, inhibitory interneurons; ◦, excitatory connection; •, inhibitory connection.

3.3 Unsupervised Learning in the Cerebral Cortex

Anatomy: A remarkable anatomical feature of the cerebral neocortex is the layered organization and the massive recurrent connections (Figure 5; see Douglas and Martin (1998) for details). The cerebral cortex is subdivided into many functional areas, where each area represents the sensory, motor, or contextual information in different modalities and frames of references (van Essen and Maunsell, 1983; Andersen, 1995; Graziano et al., 1997).
Physiology: Neurons in the cortex are well characterized by their specific response tuning, for example, to the orientation and color of the visual stimuli in the visual cortex. The response tuning of the cortical neurons is highly dependent on the sensory experience (Hirsch and Spinelli, 1970; Blakemore and Cooper, 1970). The cortical synapses have been shown to follow a Hebbian-type plasticity rule: they are potentiated when the presynaptic input is associated with the postsynaptic response (e.g., calcium ion influx) but depressed when it is not associated with postsynaptic response (Tsumoto and Suda, 1979; Artola et al., 1990).
Theoretical models: These properties strongly suggest that the information coding in the cerebral cortical areas are established by the unsupervised learning paradigm in which the activity is determined by a relaxation dynamics (e.g., as in (17)) and the synapses are update by a Hebbian rule (e.g., as in (18)).

There have been many learning and development models of response tuning in cortical neurons based on Hebbian synaptic mechanisms (von der Malsburg, 1973; Amari and Takeuchi, 1978; Grossberg, 1976; Fukushima, 1980; Kohonen, 1982). Other models have used information theoretic formulation of unsupervised learning (Linsker, 1986; Sanger, 1989; Bell and Sejnowski, 1997; Olshausen and Field, 1996).

4 Global Architectures for Adaptive Behaviors

The neural circuits of invertebrates can perform amazingly sophisticated functions despite their small size (Stein et al., 1997), since they are highly optimized for specific behaviors as a result of evolutionary processes. In contrast, the most recently developed areas of the mammalian brain, namely, the cerebellum, the basal ganglia, and the cerebral cortex, appear to be specialized for implementations of specific learning algorithms, namely, supervised, reinforcement, and unsupervised learning, rather than being genetically preprogrammed for a fixed repertoire of behaviors. Each of the three structures has modular organization: microzones in the cerebellum, output channels in the basal ganglia, and cytoarchitectonically defined areas and columnar organization in the cerebral cortex.

This section explores the possible global architectures in which the modules specialized in supervised, reinforcement, and unsupervised learning are combined into a goal-directed behaving system. Candidates of such global architectures are found in the theories of dynamic programming (Bertsekas and Tsitsiklis, 1996), reinforcement learning (Sutton and Barto, 1998), adaptive control (Narendra and Parthasarathy, 1990), and adaptive signal processing (Haykin, 1996). It is also considered whether such global architectures can be implemented under the constraints of brain anatomy.

The general construction principles are the following.

1. **Supervised learning** modules in the cerebellum can be used as the *internal model* of the environment (Kawato and Gomi, 1992b; Wolpert et al., 1998) and short-cut models of input-output mappings that have been acquired elsewhere in the brain (Ito, 1993).

2. **Reinforcement learning** modules in the basal ganglia are used to evaluate the given state and to select an action based on the evaluation.

3. **Unsupervised learning** modules in the cerebral cortex work as the medium for representing the state of the external environment or the internal context. They also provide the common representational basis for the cerebellum and the basal ganglia, between which there are no direct anatomical connections.

4.1 Reactive Action Selection

The most primitive adaptive action selection is implemented by the *actor-critic* architecture (Barto et al., 1983) illustrated in Figure 6. A stochastic policy G is implemented in the *actor* network (13). The *critic* monitors the state of the environment and estimates the value function $V(\mathbf{x})$ for the current policy.

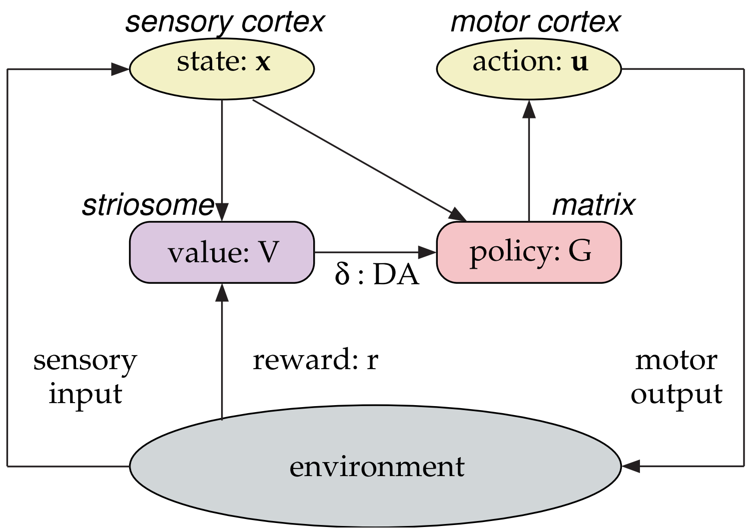

Figure 6: Model-free, stochastic action selection.

This basic reinforcement learning architecture can be implemented in the network linking frontal cortical areas, such as the supplementary motor area, and the basal ganglia, with the striatal *striosome* compartment implementing the critic and the *matrix* compartment implementing the actor (Barto et al., 1995; Houk et al., 1995).

It is essential that an appropriate encoding of the state $\mathbf{x}$ and the action $\mathbf{u}$ be available for such a simple reinforcement learning architecture to work efficiently. The cerebral cortical areas which send the input to and receive the output from the basal ganglia can subserve the function of state and action encoding based on the unsupervised learning paradigm.

4.2 Predictive Action Selection

Another action selection strategy based on the value function is to use it in conjunction with a model of the environmental dynamics. According to the theory of dynamic programming, the optimal action can be taken from the value function that satisfies the *Bellman equation*

$$V(\mathbf{x}(t)) = \max_{\mathbf{u}}(t)[r(t+1) + \gamma V(\mathbf{x}(t+1))]. \tag{19}$$

Here, the immediate future reward is predicted by a reward model

$$r(t+1) = R(\mathbf{x}(t), \mathbf{u}(t)) \tag{20}$$

and the future state $\mathbf{x}(t+1)$ is predicted by the environmental model

$$\mathbf{x}(t+1) = F(\mathbf{x}(t), \mathbf{u}(t)). \tag{21}$$

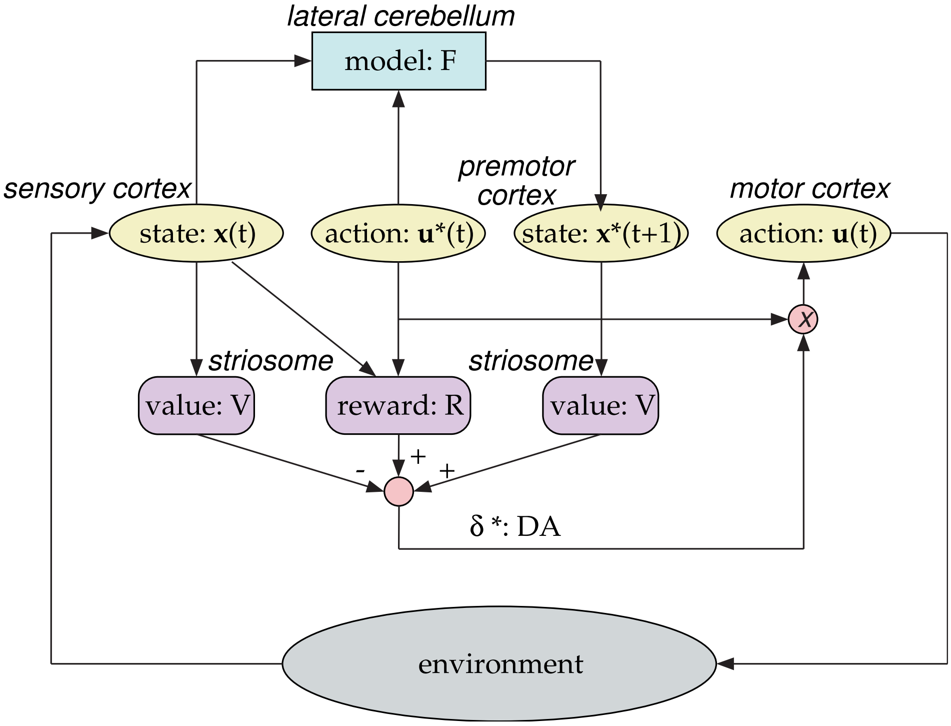

Figure 7: Action selection with a forward model.

Given such a value function V, the best action can be selected as the one that maximizes the right-hand side of the Bellman equation (19), i.e., the action that maximizes the predicted sum of immediate and future rewards.

4.2.1 Discrete Model-based Action Selection

In the tasks that involve the selection of discrete actions, it is possible to compare the right-hand side of (19) for each candidate action and then execute the action that maximizes it. One implementation of such an action selection is to consider a candidate action $u^*(t)$ one at a time, predict the resulting future state $\mathbf{x}^*(t+1)$ and its value $V(\mathbf{x}^*(t+1))$, and accept it for execution if it is good enough. Figure 7 illustrates a simple implementation of this serial action selection mechanism. In this case, whether the candidate action is good enough for execution is determined by

$$\delta^*(t+1) = R(\mathbf{x}(t), \mathbf{u}(t)) + \gamma V(\mathbf{x}^*(t+1)) - V(\mathbf{x}(t)), \tag{22}$$

which has the same form as the TD error (10).

It is possible for such an architecture to be implemented in the network linking the prefrontal and rostral premotor areas, the lateral cerebellum, and the basal ganglia as shown in Figure 7. The prefrontal and the rostral premotor areas as well as the lateral cerebellum

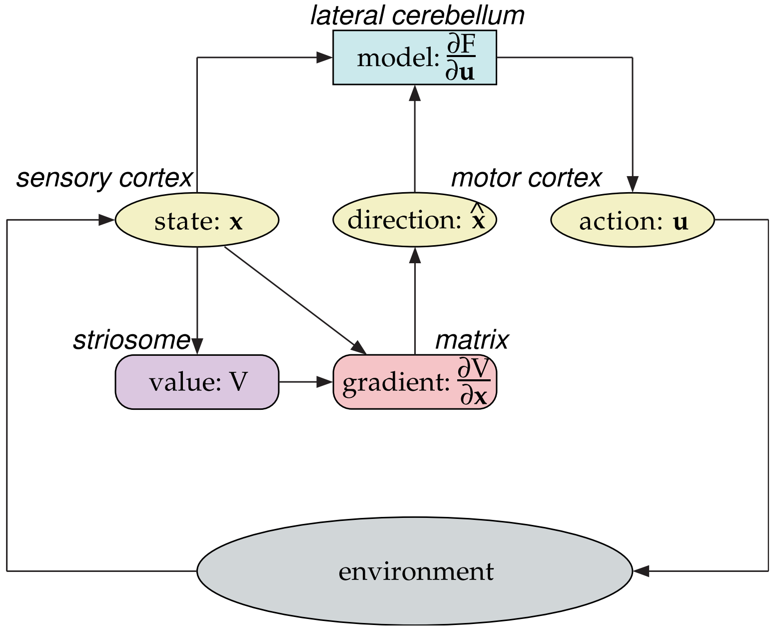

Figure 8: Differential model-based action selection.

are known to be activated during imagery of body movement and during mental operations such as mathematical calculation (Deiber et al., 1998; Rueckert et al., 1996).

Thus, the loop consisting of the lateral cerebellum and the prefrontal and rostral premotor cortex should be capable of predicting the outcome of a hypothetical action. These cortical areas also send projection to the anterior part of the basal ganglia, where the multiple inhibitory pathways should enable the computation of the predictive TD error (22) encoded in the activity of the dopamine neurons. In fact, the anterior basal ganglia, the pre-SMA, and the prefrontal cortex have been shown to be preferentially involved in learning a new sequential movement by trial and error in both animal lesion studies (Miyachi et al., 1997) and human imaging studies (Hikosaka et al., 1995; Sakai et al., 1998).

4.2.2 Differential model-based action selection

When continuous actions are considered, action selection based on the Bellman equation (19) is achieved by differentiating the maximization term for action $\mathbf{u}$ and making it zero, namely

$$\frac{\partial R(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}} + \gamma \frac{\partial V(\mathbf{x})}{\partial \mathbf{x}} \frac{\partial F(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}} = 0. \tag{23}$$

By assuming that the reward (cost) for the action $\mathbf{u}$ is quadratic, the term $\frac{\partial R(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}}$ becomes a linear function of $\mathbf{u}$ and thus the optimal action is given by

$$\mathbf{u} \propto \frac{\partial F(\mathbf{x},\mathbf{u})'}{\partial \mathbf{u}} \frac{\partial V(\mathbf{x})'}{\partial \mathbf{x}}. \tag{24}$$

In this equation, $\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}'$ represents the steepest ascent direction of the value function, which can be regarded as the "desired" movement direction. The input gain matrix $\frac{\partial F(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}}'$, which can be regarded as a local inverse dynamic model, converts this movement direction into action direction (Doya, 1996). The corresponding network architecture is illustrated in Figure 8.

Such a control architecture could possibly be implemented in the global network linking the basal ganglia, the primary motor cortex (M1), and the lateral cerebellum, as shown in Figure 8. M1 is reciprocally connected to both the cerebellum and the basal ganglia. Georgopoulos (1995) and Scott (1997) have observed cellular responses encoding both movement direction and muscular activities.

The input gain $\frac{\partial F(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}}'$ could possibly be implemented in the lateral cerebellum by supervised learning of state change $\dot{\mathbf{x}}$ for action $\mathbf{u}(t)$. Although there is support for the hypothesis that the value function $V(\mathbf{x})$ is learned in the basal ganglia (Barto et al., 1995), it is not certain whether the gradient $\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}$ of the value function can be calculated in the basal ganglia. However, it may be possible to learn to approximate such a quantity by way of the temporal differential formula $\dot{V} = \frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}\dot{\mathbf{x}}$.

This architecture shares the same features with the previously proposed model of the limb premotor network (Houk et al., 1993), in that there is no explicit desired trajectory calculation and the detailed motor command is generated in the reciprocal connection between the motor cortex and the cerebellum.

4.3 Models for Prediction, Simulation, and Encapsulation

Let us now consider several ways of utilizing cerebellar supervised learning modules in conjunction with any of the above action selection architectures.

4.3.1 Model-based state estimation

One big challenge in the control of movement is the existence of feedback delays and hidden states. The basic assumption in the above action architectures is that the information necessary for evaluation and action selection is readily available in the form of the state vector $\mathbf{x}(t)$. One possible role of the cerebellum and the cerebral cortex is to provide the estimate of the current state despite the delayed, noisy, and/or degenerate sensory inputs.

This could possibly be implemented by the network dynamics within the cortical areas or by the network linking the cerebral cortex and the cerebellum, which is supposed to provide the model of the environmental dynamics (Wolpert et al., 1998), as shown in Figure 9. Typical examples of state estimation using dynamic models are the Smith predictor (Miall et al., 1993) and the Kalman filter (Paulin, 1993).

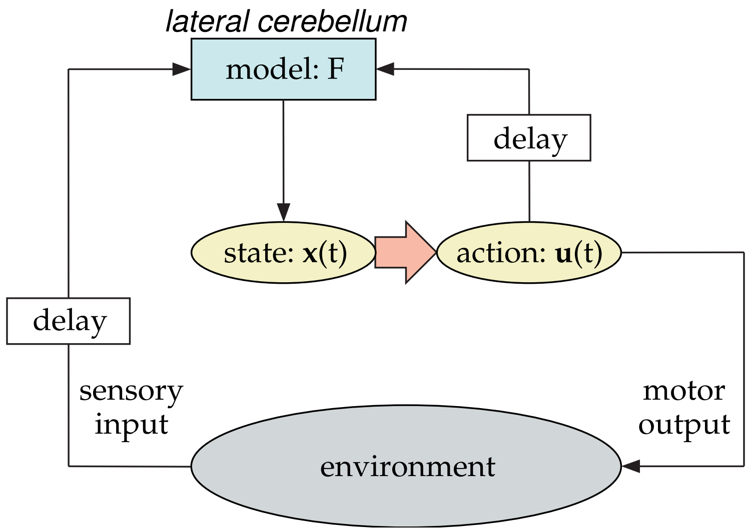

Figure 9: Compensation of sensory feedback delay with a forward model. The thick arrow represents either of the state-to-action mapping by the architectures shown in Figures 6, 7, and 8.

Such model-based estimation of the external state is useful not only for motor execution, but also for sensory acquisition that involves a comparison of actual and expected sensory outcome, for example, somatosensory object recognition (Gao et al., 1996) and judgment of precise timing (Ivry, 1996).

4.3.2 Simulation in the virtual environment

In the model-based, predictive action selection architecture above, we only considered a single-step planning of action. However, provided that the number of possible actions are small and that there is enough working memory capacity and time for decision making, it is possible to extend such mechanisms to multiple-step prediction and evaluation.

It is also possible to use such an internal model of environment not only for on-line action selection but also for the improvement of the value function and the policy based on the simulated experiences, as illustrated in Figure 10. Such *off-line* learning could enable learning a behavior from a very limited number of actual experiences, although its success depends strongly on the accuracy of the environmental model.

4.3.3 Encapsulating learned mappings

Many of the discussed control and decision architectures involve global communication between the cerebellum, the basal ganglia, and the cerebral cortex, and this can be quite costly in the hardware resources and processing time. These architectures are useful for

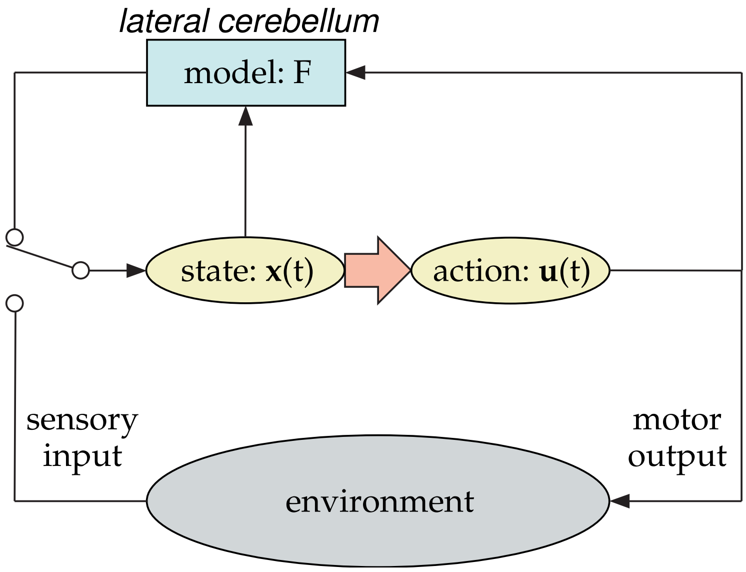

Figure 10: Simulation of behavior using a forward model instead of the real environment.

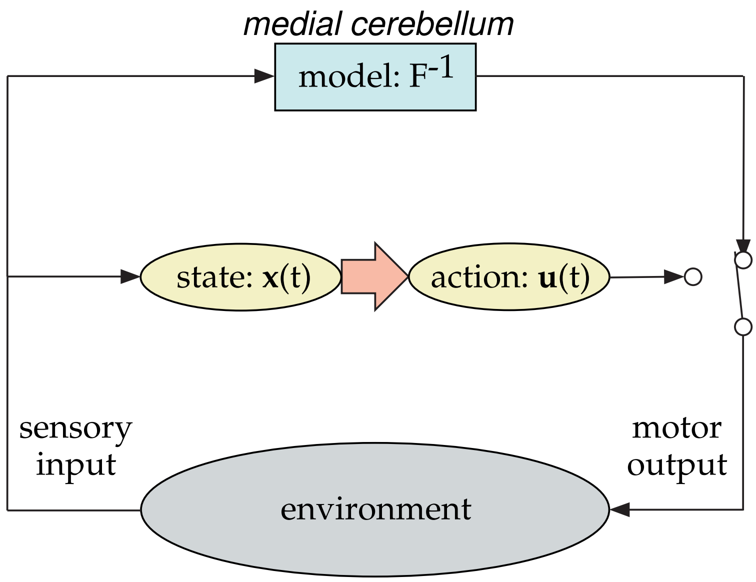

Figure 11: Encapsulation of complex decision process in a simple reactive mapping.

learning a new task, but once it has been well learned, it is more advantageous to store the learned policy in a local circuit for compact storage, reliable execution, and quick reaction.

For example, supervised learning modules in the medial part of the cerebellum, which receives the spinal sensory afferents and sends out motor commands to the spinal cord and primary motor cortex, are a candidate for a repository of learned sensory-motor mapping (Figure 11). For example, in arm movement control, the output the cortical feedback pathway could be used as the teacher signal for the cerebellum. In this case, the cerebellum serves as a direct controller (Barto et al., 1996), or an inverse model of the arm dynamics (Kawato et al., 1987; Schweighofer et al., 1998).

Similarly, supervised learning modules in the lateral part of the cerebellum, which are reciprocally connected to different parts of the frontal cortex including the prefrontal cortex, are a candidate for the repository of more abstract state-action mapping that was originally acquired using the global network, as illustrated in Figure 7.

Such encapsulation of learned mapping is useful for robust storage and quick reaction, and it frees a large portion of the brain from the task already learned so that these portions of the brain can be used to learn new tasks.

5 Brain Activation in Sequence Learning

The characterization of specific computations for the cerebellum, the basal ganglia, and the cerebral cortex and their possible combinations as a global learning architecture might be useful in inferring the specific roles of the activated areas in brain imaging studies. The results of functional imaging experiments on sequence learning are investigated in view of the architectures listed above.

5.1 Implicit Learning

In an implicit sequence learning task paradigm, subjects were asked to perform simple actions (usually press buttons), initially by following cues shown in random order. Then, the order of cues (thus the actions) was fixed in a particular sequence or generated according to a certain grammar. As a subject repeated the regular sequence, the reaction time for pressing buttons became shorter even when the subject was not aware of any particular order or grammar.

Table 1A summarizes the activation in the cerebellum, the basal ganglia, and the frontal cortex during such implicit sequence learning. The most consistent activity was found in the basal ganglia. In view of the global architectures, the result is compatible with the reactive architecture shown in Figure 6, which does not use an explicit model of the environmental dynamics, in this case the button sequence.

Table 1: Activations of prefrontal cortex, premotor cortex, basal ganglia, and cerebellum in sequence learning tasks. Abbreviations: r., right; l., left; b., bilateral; a., anterior; v., ventral., DLPF, dorsolateral prefrontal cortex; PM, premotor area; SMA, supplementary motor area, GP, global pallidus.

	Prefrontal	Premotor	Basal ganglia	Cerebellum
A. Implicit Learning				
Grafton et al. (1995)		b.SMA	b.putamen	r.dentate
Rauch et al. (1995)		r.PM	r.striatum	
Doyon et al. (1996)			r.v.striatum	r.dentate
Berns et al. (1997)		l.PM	r.v.striatum	
Grafton et al. (1998)		r.SMA		
B. Exploratory Learning				
Jenkins et al. (1994)	b.	b.PM		b.cortex b.nuclei
Hikosaka et al. (1996)		preSMA		
Jeuptner et al. (1997)	b.	b.PM	b.caudate b.GP	b.cortex b.nuclei
Sakai et al. (1998)	l.DLPF	preSMA		

5.2 Sequence Learning by Exploration

Let us now focus on sequence learning by trial and error. Subjects were asked to learn the right sequential movements (usually pressing buttons with fingers) by actively searching for the right order. Although such a task can be learned with the simple stochastic learning architecture shown in Figure 6, more efficiently learning is possible with the predictive model of the task, as shown in Figure 7. For example, it is possible to avoid making the same errors with the ability to predict the outcome of the action taken previously. Such predictive action selection is particularly useful early in the learning process before a stereotyped reaction is established.

Table 1B summarizes the areas activated during such exploratory sequence learning. Prominent activations were found in the prefrontal cortex and the rostral premotor areas. This is compatible with the predictive action selection architecture shown in Figure 7, which relies on the state prediction presented in the prefrontal and premotor cortex.

6 Discussion

This paper presented a hypotheses in which the cerebellum, the basal ganglia, and the cerebral cortex are, respectively, specialized in supervised, reinforcement, and unsupervised learning paradigms. First, the anatomical, physiological, and theoretical support for the hypothesis was outlined. Then, a number of global architectures for goal-oriented behaviors were considered with the combination of those learning modules within biological constraints.

6.1 Previous Models

Despite the wealth of the models for the cerebellum, the basal ganglia, and the cerebral cortex, there have been few attempts at characterizing the specific computations in these three brain areas in a unified framework.

Thach et al. (1993) compared the roles the cerebellum and the basal ganglia and concluded that the cerebellum is involved in combining primitive movements and that the basal ganglia is involved in the inhibition and disinhibition of primitive or complex movements. Mushiake (1995) gave a nice review of anatomical and physiological findings that point to different functions of the cerebellum, the basal ganglia, and the cerebral cortex.

Houk and Wise (1995) emphasized the distributed nature of the operation of the networks liking the basal ganglia, the cerebellum and the cerebral cortex and characterized their major functions as "context recognition" for the basal ganglia, "pattern generation" for the cerebellum, and "information stores" for the cerebral cortex. However, they mentioned differences in the nature of computations for the evidence in support of their ideas: supervised learning in the cerebellum, reinforcement learning in the basal ganglia, and attractor dynamics in the cerebral cortex.

The novel features in the current proposal are: 1) the differences in the computational paradigms are the primary differences in the three structures, and 2) the theory is based on sound mathematical frameworks, including dynamic programming, reinforcement learning, and adaptive control.

The first point enables a very simple and coherent characterization of the three structures. The second point enables a systematic exploration of the possible ways of their interactions. These models are useful both in interpreting the existing experimental data and in planning new behavioral experiments. Furthermore, it is possible to test the performance of those architectures in a realistic simulation of experimental paradigms.

One such example is seen in the model of sequence learning by Nakahara et al., (1998). They succeeded in replicating the experimental results from a "2x5 task" based on the assumption that the loops connecting the basal ganglia and the cerebral cortex work as the "actor-critic" architecture (Figure 6) and that different cortical areas provide the sensory and contextual information in different frames of reference.

6.2 Cerebral and Cerebellar Mappings

One of the major functions of the cerebellum is to provide a model or mapping between sensory-motor representations. However, it is well known that different cortical areas have different maps that capture different aspects of the sensory and motor state, for example, a visuo-motor target in different frames of references (Snyder et al., 1998; Graziano et al., 1997). Thus, the connection between different cortical areas can work as a *coordinate transformation* system. What are the essential difference between cortical and cerebellar mapping?

As described in Section 2, unsupervised learning is based on the *statistical* properties of the signals and a solution depends on iterative network dynamics and is quite context dependent. On the other hand, a deterministic function of the current input is learned in the feed-forward cerebellar circuit. Thus, a cerebellar mapping can work as a short-cut circuit or a look-up table for a mapping that was originally developed by the time-consuming cortico-cortical processing. Such a mechanism is especially useful when the same mapping is used repeatedly for time critical tasks, for example, inverse kinematic and inverse dynamic mappings for arm movement.

6.3 Cognitive Tasks

In this paper, the considerations were mainly made in the context of motor control, but similar architectures could also be utilized for cognitive functions including logical inference and communication. For example, in mathematical problem-solving, sequential selection of arithmetic operations on symbolic representations can be implemented by the predictive decision architecture shown in Figure 7. After repeated problem solving, the search process can be skipped by using the problem-solution mapping learned in the cerebellum, as shown in Figure 11.

In the context of communication, the "environment" is the partner of communication and the goal is to bring the physical or internal state of the partner into a desired state. This involves sequential selection of actions, i.e., words or gestures, in an appropriate sequence, in the same way as in the case of many control tasks. When the model of the partner is available, the goal can be achieved more readily and quickly. If the internal models of the speaker and listener are similar, communication is made efficient. efficient.

Actually, the brain areas that are activated in complex motor tasks are also activated in cognitive or verbal tasks (Mellet et al., 1996; Rueckert et al., 1996). Thus, we postulate that the current investigation of behavioral architectures and their brain implementation is not only applicable to motor tasks but also to cognitive tasks including communication.

Predictive action selection architecture (Figure 7) has been extensively used in traditional artificial intelligence (AI) research. Classical AI approaches to an understanding of human cognitive functions failed because these approaches usually took into account only a single or a few possible architectures for problem solving. The massively parallel organization of the brain enables the use of many different strategies for the same task. It is essential to consider the process in which a deliberate, intensive task turns into a trivial, automatic task after repeated training.

6.4 Who Decides the Architecture?

This paper has presented useful module combinations but has not considered "how" such appropriate modules are picked up from the cerebellum, the basal ganglia, and the cerebral cortex and are assembled into a global network. The genetic information of area to area connections is certainly important, but it does not explain how the exact networks for novel, learned behaviors is set up.

Traditionally, the *central executive* was supposed to undertake the task of activating and connecting appropriate processing modules (Baddeley, 1992). However, its computational principle and biological substrate is unknown. Proponents of *subsumption architectures* would reject the need for such a centralized process. They have demonstrated the recruitment of appropriate modules by simple rules (Brooks, 1991). However, the way different modules are activated and preceded is usually pre-wired in many robotic demonstrations.

An alternative possibility is that the competitive learning mechanism in the cerebral cortex is utilized for the selection of appropriate modules for the environmental state and the goal of the task (Houk and Wise, 1995). Recently, a computational framework of selecting and combining multiple control modules based on the performances of the multiple predictive models has been proposed (Wolpert and Kawato, 1998; Wolpert et al., 1998). Such a modular architecture could possibly be implemented by a network linking the cerebellum and the prefrontal and rostral premotor cortex. An activation of the preSMA and the rostral cingulate motor area (CMAr) at the time of change in the motor plan is in accordance with such a possibility (Shima et al., 1996; Shima and Tanji, 1998).

6.5 Who Sets the Parameters?

The performance of any learning algorithm depends on the setting of some global parameters, such as the learning rates. The appropriate values of such parameters are usually task dependent and therefore have to be under active regulation. In the framework of reinforcement learning, such global parameters include the time scale of an evaluation (discount factor), the size of the noise for exploration (temperature), the magnitude of cost for action, and the learning rates. How does the brain appropriately set such parameters?

This is still an open problem, but the likely candidates are the diffuse ascending projections from the midbrain nuclei to the cerebellum, the basal ganglia, and the cerebral cortex. For example, the noradrenergic system is known to modulate the sharpness of neural responses (Usher et al., 1999), which is similar to the role of temperature in many learning algorithms. The level of serotonin affects the subject's activeness, which is reminiscent of the role of the action cost parameter. The cholinergic and noradrenergic input is known to regulate the synaptic plasticity of the cortical neurons (Brocher et al., 1992). Computational studies of regulatory mechanisms for learning systems, or meta-learning, would provide a theoretical basis for understanding the roles of emotional systems in regulating behavioral learning.

Acknowledgments

I thank Mitsuo Kawato, Takashi Omori, and two anonymous reviewers for their comments on the manuscript. I am grateful to Minoru Kimura, James Houk, Okihide Hikosaka, Hiro Nakahara, Ann Graybiel, Jun Tanji, and Manabu Honda for their helpful discussions.

References

Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Bioscience*, 10:25–61.

Alexander, G. E. and Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neuroscience*, 13:266–271.

Amari, S. and Takeuchi, A. (1978). Mathematical theory on formation of category detecting nerve cells. *Biological Cybernetics*, 29:127–136.

Andersen, R. A. (1995). Encoding of intention and spatial location in the posterior parietal cortex. *Cerebral Cortex*, 5:457–469.

Aosaki, T., Tsubokawa, H., Watanabe, K., Graybiel, A. M., and Kimura, M. (1994). Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensory-motor conditioning. *Journal of Neuroscience*, 14:3969–3984.

Artola, A., Brocher, S., and Singer, W. (1990). Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature*, 347:69–72.

Baddeley, A. D. (1992). Working memory. *Science*, 255:556–559.

Barto, A. G. (1995). Adaptive critics and the basal ganglia. In Houk, J. C., Davis, J. L., and Beiser, D. G., editors, *Models of Information Processing in the Basal Ganglia*, pages 215–232. MIT Press, Cambridge, MA, USA.

Barto, A. G., Bradtke, S. J., and Singh, S. P. (1995). Learning to act using real-time dynamic programming. *Artificial Intelligence*, 72:81–138.

Barto, A. G., Buckingham, J. T., and Houk, J. C. (1996). A predictive switching model of cerebellar movement control. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., editors, *Advances in Neural Information Processing Systems 8*, pages 138–144. MIT Press, Cambridge, MA.

Barto, A. G., Sutton, R. S., and Anderson, C. W. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:834–846.

Becker, S. (1995). Unsupervised learning with global objective functions. In Arbib, M. A., editor, *The Handbook of Brain Theory and Neural Networks*, pages 997–1000. MIT Press, Cambridge, MA, USA.

Bell, A. J. and Sejnowski, T. J. (1997). The independent components of natural scenes are edge filters. *Vision Research*, 37:3327–3338.

Berns, G., Cohen, J. D., and Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science*, 276:1272–1275.

Berns, G. S. and Sejnowski, T. J. (1998). A computational model of how the basal ganglia produces sequences. *Journal of Cognitive Neuroscience*, 10:108–121.

Bertsekas, D. P. and Tsitsiklis, J. N. (1996). *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA, USA.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, UK.

Blakemore, C. and Cooper, G. F. (1970). Development of the brain depends on the visual environment. *Nature*, 228:477–478.

Bower, J. M. (1997). Is the cerebellum sensory for motor's sake, or motor for sensory's sake: the view from the whisker of a rat? *Progress in Brain Research*, 114:463–96.

Brocher, S., Artola, A., and Singer, W. (1992). Agonists of cholinergic and noradrenergic receptors facilitate synergistically the induction of long-term potentiation in slices of rat visual cortex. *Brain Research*, 573:27–36.

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

Brown, L. L., Schneider, J. S., and Lidsky, T. I. (1997). Sensory and cognitive functions of the basal ganglia. *Current Opinion in Neurobiology*, 7:157–163.

Calabresi, P., Pisani, A., Mercuri, N. B., and Bernardi, G. (1996). The corticostriatal projection: from synaptic plasticity to dysfunctions of the basal ganglia. *Trends in Neurosciences*, 19:19–24.

Daniel, H., Levenes, C., and Crépel, F. (1998). Cellular mechanisms of cerebellar ltd. *Trends in Cognitive Sciences*, 2:401–407.

Deiber, M. P., Ibanez, V., Honda, M., Sadato, N., Raman, R., and Hallett, M. (1998). Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography. *Neuroimage*, 7:73–85.

Desmond, J. E. and Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: language, learning and memory. *Trends in Cognitive Sciences*, 2:355–362.

DeZeeuw, C. I., Simpson, J. I., Hoogenraad, C. C., Galjart, N., Koekkoek, S. K. E., and Ruigrok, T. J. H. (1998). Microcircuitry and function of the inferior olive. *Trends in Cognitive Sciences*, 2:391–400.

Dominey, P., Arbib, M., and Joseph, J.-P. (1995). A model of corticostriatal plasticity for learning oculomotor associations and sequences. *Journal of Cognitive Neurosciecen*, 7:311–336.

Douglas, R. and Martin, K. (1998). Neocortex. In Shepherd, G. M., editor, *The Synaptic Organization of the Brain*, pages 459–509. Oxford University Press, New York, 4th edition edition.

Doya, K. (1996). Temporal difference learning in continuous time and space. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., editors, *Advances in Neural Information Processing Systems 8*, pages 1073–1079. MIT Press.

Doyon, J., Owen, A. M., Petrides, M., Sziklas, B., and Evans, A. C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *European Journal of Neuroscience*, 8:637–648.

Fujita, M. (1982). Adaptive filter model of the cerebellum. *Biological Cybernetics*, 45:195–206.

Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36:193–202.

Gao, J. H., Parsons, L. M., Bower, J. M., Xiong, J., Li, J., and Fox, P. T. (1996). Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science*, 272:545–547.

Georgopoulos, A. P. (1995). Current issues in directional motor control. *Trends in Neurosciences*, 18:506–510.

Gomi, H. and Kawato, M. (1992). Adaptive feedback control models of the vestibulocerebellum and spinocerebellum. *Biological Cybernetics*, 68:105–114.

Gomi, H., Shidara, M., Takemura, A., Inoue, Y., Kawano, K., and Kawato, M. (1998). Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys. I. Simple spikes. *Journal of Neurophysiology*, 80:818–831.

Grafton, S. T., Hazeltine, E., and Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience*, 7:497–510.

Grafton, S. T., Hazeltine, E., and Ivry, R. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *Journal of Neuroscience*, 18:9420–9428.

Graybiel, A. M. (1995). Building action repertoires: memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology*, 5:733–741.

Graziano, M. S. A., Hu, X. T., and Gross, C. G. (1997). Visuo-spatial properties of ventral premotor cortex. *Journal of Neurophysiology*, 77:2268–2292.

Grossberg, S. (1976). On the development of detectors in the visual cortex with applications to learning and reaction-diffusion systems. *Biological Cybernetics*, 21:145–159.

Harpur, G. F. and Prager, R. W. (1995). Techniques in low entropy coding with neural networks. CUED/F-INFENG/TR 197, Cambridge University Engineering Department, Cambridge, England.

Haykin, S. (1996). *Adaptive Filter Theory.* Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition edition.

Hikosaka, O., Miyachi, S., Miyashita, K., and Karadi, Z. (1995). Possible roles of the basal ganglia in procedural learning. In *Fourth IBRO World Congress of Neuroscience Abstracts*, page 55, Oxford. IBRO, Rapid Communications.

Hikosaka, O., Sakai, K., Miyauchi, S., Takino, R., Sasaki, Y., and Pütz, B. (1996). Activation of human presupplementary motor area in learning of sequential procedures: A functional MRI study. *Journal of Neurophysiology*, 76:617–621.

Hirsch, H. V. B. and Spinelli, D. N. (1970). Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cat. *Science*, 168:869–871.

Houk, J. C. (1997). On the role of the cerebellum and basal ganglia in cognitive signal processing. *Progress in Brain Research*, 114:545–554.

Houk, J. C., Adams, J. L., and Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In Houk, J. C., Davis, J. L., and Beiser, D. G., editors, *Models of Information Processing in the Basal Ganglia*, pages 249–270. MIT Press, Cambridge, MA.

Houk, J. C. and Alford, S. (1996). Computational significance of the cellular mechanisms for synaptic plasticity in Purkinje cells. *Behavioral and Brain Sciences*, 19:457–461.

Houk, J. C., Buckingham, J. T., and Barto, A. G. (1996). Models of the cerebellum and motor learning. *Behavioral and Brain Sciences*, 19:368–383.

Houk, J. C., Keifer, J., and Barto, A. G. (1993). Distributed motor commands in the limb premotor networks. *Trends in Neuroscience*, 16:27–33.

Houk, J. C. and Wise, S. P. (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: Their role in planning and controlling action. *Cerebral Cortex*, 2:95–110.

Ito, M. (1984). *The Cerebellum and Neural Control*. Raven Press, New York.

Ito, M. (1993). Movement and thought: identical control mechanisms by the cerebellum. *Trends in Neurosciences*, 16:448–450.

Ito, M., Sakurai, M., and Tongroach, P. (1982). Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar purkinje cells. *Journal of Physiology*, 324:113–134.

Ivry, R. B. (1996). The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology*, 6:851–857.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., and Passingham, R. E. (1994). Motor sequence learning: a study with positron emission tomography. *Journal of Neuroscience*, 14:3775–3790.

Jeuptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. J., and Passingham, R. E. (1997). Anatomy of motor learning. I. Frontal cortex and attention to action. *Journal of Neurophysiology*, 77:1313–1324.

Kawato, M., Furukawa, K., and Suzuki, R. (1987). A hierarchical neural network model for control and learning of voluntary movement. *Biological Cybernetics*, 57:169–185.

Kawato, M. and Gomi, H. (1992a). The cerebellum and VOR/OKR learning models. *Trends in Neurosciences*, 15:445–453.

Kawato, M. and Gomi, H. (1992b). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological Cybernetics*, 68:95–103.

Kettner, R. E., Mahamud, S., Leung, H.-C., Sitkoff, N., and Houk, J. C. (1997). Prediction of complex two-dimensional trajectories by a cerebellar model of smooth pursuit eye movement. *Journal of Neurophysiology*, 77:2115–2130.

Kitazawa, S., Kimura, T., and Yin, P.-B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature*, 392:494–497.

Kobayashi, Y., Kawano, K., Takemura, A., Inoue, Y., Kitama, T., Gomi, H., and Kawato, M. (1998). Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys. II. Complex spikes. *Journal of Neurophysiology*, 80:832–848.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69.

Lawrence, A. D., Sahakian, B. J., and Robbins, T. W. (1998). Cognitive functions and corticostriatal circuits: insights from huntington's disease. *Trends in Cognitive Sciences*, 2:379–388.

Leiner, J. C., Leiner, A. L., and Dow, R. S. (1993). Cognitive and language functions of the cerebellum. *Trends in Neurosciences*, 16:444–447.

Linsker, R. (1986). From basic network principles to neural architecture. *Proceedings of the National Academy of Sciences, USA*, 83:7508–7512, 8390–8394, 8779–8783.

Llinas, R. R. and Walton, K. D. (1998). Cerebellum. In Shepherd, G. M., editor, *The Synaptic Organization of the Brain*, pages 255–288. Oxford University Press, New York, 4th edition edition.

Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology*, 202:437–470.

Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., and Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *Journal of Neuroscience*, 16:6504–6512.

Miall, R. C., Weir, D. J., Wolpert, D. M., and Stein, J. F. (1993). Is the cerebellum a Smith predictor? *Journal of Motor Behavior*, 25:203–216.

Middleton, F. A. and Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science*, 266:458–461.

Middleton, F. A. and Strick, P. L. (1996). New concepts regarding the organization of basal ganglia and cerebellar output. In Ito, M. and Miyashita, Y., editors, *Integrative and Molecular Approach to Brain Function*, pages 253–269. Elsevier Science.

Middleton, F. A. and Strick, P. L. (1998). Cerebellar output: motor and cognitive channels. *Trends in Cognitive Sciences*, 2:348–354.

Miyachi, S., Hikosaka, O., Miyashita, K., Karadi, Z., and Rand, M. K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. *Experimental Brain Research*, 151:1–5.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, 16:1936–1947.

Mushiake, H. (1995). Functional relationship of the basal ganglia and cerebellum with motor cortices. *Advances in Neurological Sciences*, 39:277–289. (in Japanese).

Nakahara, H., Doya, K., Hikosaka, O., and Nagano, S. (1997). Multiple representations in the basal ganglia loops for acquisition and execution of sequential motor control. In *Society for Neuroscience Abstracts*, volume 23, page 778.

Nakahara, H., Doya, K., Hikosaka, O., and Nagano, S. (1998). Reinforcement learning with multiple representations in the basal ganglia loops for sequential motor control. In *International Joint Conference on Neural Networks*, pages 1553–1558.

Narendra, K. and Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, 1:4–27.

Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning sparse code for natural images. *Nature*, 381:607–609.

Paulin, M. G. (1993). The role of the cerebellum in motor control and perception. *Brain, Behavior, and Evolution*, 41:39–50.

Rauch, S. L., Savage, C. R., Brown, H. D., Curran, T., Alpert, N. M., Kendrick, A., Fischman, A. J., and Kosslyn, S. M. (1995). A PET investigation of implicit and explicit sequence learning. *Human Brain Mapping*, 3:271–286.

Raymond, J. L., Lisberger, S. G., and Mauk, M. D. (1996). The cerebellum: a neuronal learning machine? *Science*, 272:1126–1131.

Rueckert, L., Lange, N., Partiot, A., Appollonio, I., Litvan, I., Le Bihan, D., and Grafman, J. (1996). Visualizing cortical activation during mental calculation with functional MRI. *Neuroimage*, 3:97–103.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., and Putz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor sequence learning. *Journal of Neuroscience*, 18:1827–1840.

Sanger, T. D. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2:459–473.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80:1–27.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, 13:900–913.

Schultz, W., Apicella, P., Romo, R., and Scarnati, E. (1995). Context-dependent activity in primate striatum reflecting past and future behavioral events. In Houk, J. C., Davis, J. L., and Beiser, D. G., editors, *Models of Information Processing in the Basal Ganglia*, pages 11–27. MIT Press, Cambridge, MA.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Nature*, 275:1593–1599.

Schweighofer, N., Arbib, M. A., and Dominey, P. F. (1996). A model of the cerebellum in adaptive control of saccadic gain. I. The model and its biological substrate. *Biological Cybernetics*, 75:19–28.

Schweighofer, N., Arbib, M. A., and Kawato, M. (1998). Role of the cerebellum in reaching movements in humans. II. A neural model of the intermediate cerebellum. *European Journal of Neuroscience*. 86-94.

Scott, S. H. (1997). Comparison of onset time and magnitude of activity for proximal arm muscles and motor cortical cells before reaching movements. *Journal of Neurophysiology*, 77:1016–1022.

Shidara, M., Kawano, K., Gomi, H., and Kawato, M. (1993). Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature*, 365:50–52.

Shima, K., Mushiake, H., Saito, N., and Tanji, J. (1996). Role for cells in the presupplementary motor area in updating motor plans. *Proceedings of National Academy of Sciences, USA*, 93:8694–8698.

Shima, K. and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*, 282:1335–1338.

Snyder, L. H., Grieve, K. L., Brotchie, P., and Andersen, R. A. (1998). Separate body- and world-referenced representation of visual space in parietal cortex. *Nature*, 394:887–891.

Stein, P. S. G., Grillner, S., Selverston, A. I., and Stuart, D. G. (1997). *Neurons, Networks, and Motor Behaviors*. MIT Press, Cambridge, MA.

Suri, R. E. and Schultz, W. (1998). Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Experimental Brain Research*, 121:350–354.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning*. MIT Press, Cambridge, MA, USA.

Thach, W. T., Mink, J. W., Goodkin, H. P., and Keating, J. G. (1993). Combining versus gating motor programs: Differential roles for cerebellum and basal ganglia? In Mano, M., Hamada, I., and DeLong, M. R., editors, *Roles of Cerebellum and Basal Ganglia in Voluntary Movement*, pages 235–245. Elsevier, Amsterdam.

Thompson, R. F. and Krupa, D. J. (1994). Organization of memory traces in the mammalian brain. *Annual Review of Neuroscience*, 17:519–549.

Tsumoto, T. and Suda, K. (1979). Cross-depression: an electrophysiological manifestation of binocular competition in developing visual cortex. *Brain Research*, 168:190–194.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., and Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science*, 283:549–554.

van Essen, D. and Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, 6:370–375.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14:85–100.

Wickens, J. R., Begg, A. J., and W., A. G. (1996). Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience*, 70:1–5.

Wilson, C. W. (1998). Basal ganglia. In Shepherd, G. M., editor, *The Synaptic Organization of the Brain*, pages 329–375. Oxford University Press, New York, 3rd edition edition.

Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11:1317–1329.

Wolpert, D. M., Miall, R. C., and Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2:338–347.