

An Integrated Cerebro-Cerebellar Model Demonstrating Associative Learning and Motor Control

Charles C. Peck, Tyler Streeter, and James Kozloski
IBM T.J. Watson Research Center
P.O. Box 218, Yorktown Heights, NY 10598

Abstract

A model of the cerebellum is proposed that explains cerebellar contributions to motor function and associative learning. This model is consistent with a wide range of biological observations and it places the cerebellum into a global, integrated network. A computational model is implemented using Wilson-Cowan style models, and the capabilities of associative learning and complex reaching behaviors, using a one armed automaton, are demonstrated. It is shown how the integrated cerebro-cerebellar model learns to modulate the cortex to produce specific, expected behaviors and other capabilities attributed to the cerebellum.

1 Introduction

The cerebellum has long attracted and intrigued modelers of neural structures and experimentalists. It is attractive for experimentation, study, and modeling because of its repetitive, well characterized circuitry, easily observable impact on behavior, and its responsiveness to various sensory stimuli. It is intriguing because it possesses roughly half the neurons in the brain and it is implicated in, but not required for, many seemingly disparate functions, such as associative learning, motor control, timing of movement initiation, movement termination, and offloading tasks from the supplementary motor area[7, 3, 1, 2]. Recently, the cerebellum has been implicated in cognitive tasks, such as word retrieval[1].

These observations raise many questions. How can a simple, repetitive circuit participate in so many seemingly disparate functions? Could a single function mediate all of these capabilities? What is its role in the global brain architecture?

This paper models the cerebellum in the context of the global brain architecture and proposes a role capable of explaining many of the disparate functions attributed to the cerebellum.

2 The Cerebellum in Context

To motivate the model, it is useful to redraw the cerebro-cerebellar circuit, as in Figure 1, to emphasize many salient characteristics. First, the motor output of

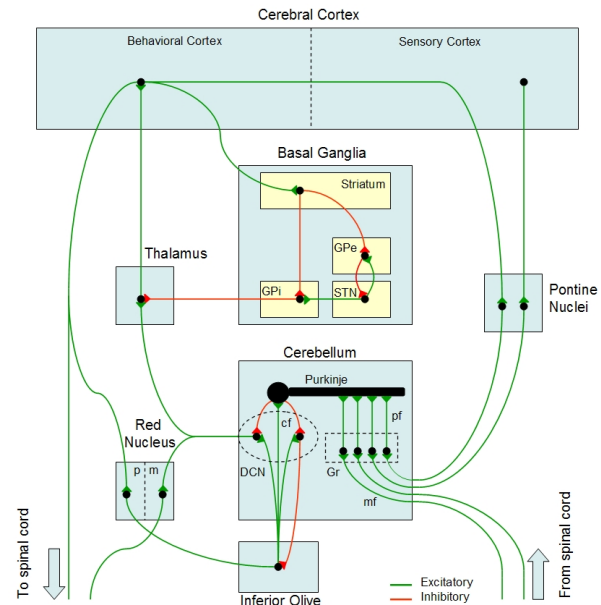


Fig. 1. Schematic of the internal circuitry of the cerebellum and the relationships between the cerebellum and other brain structures. Triangular terminations denote synaptic targets. DCN: deep cerebellar nuclei, cf: climbing fibers, pf: parallel fibers, mf: mossy fibers, Gr: granule cells, STN: subthalamic nucleus, GPi, GPe: internal and external segments of the global pallidus, p: parvocellular, m: magnocellular. Color key included above.

the brain is mediated by projections to the spinal cord emanating from the motor areas of the behavioral cortex¹ and the magnocellular neurons of the red nucleus[1]. The red nucleus, in turn, appears to act as a relay for the excitatory outputs of the deep cerebellar nuclei.

Second, there appears to be two point-to-point, registered loops involving the projection neurons in the motor areas of the behavioral cortex. The first can be traced

¹Here, the behavioral cortex refers to the frontal lobe. The motor areas are those receiving afferents from the ventral lateral nucleus of the thalamus (i.e., the primary motor, premotor, and supplementary motor areas[1]).

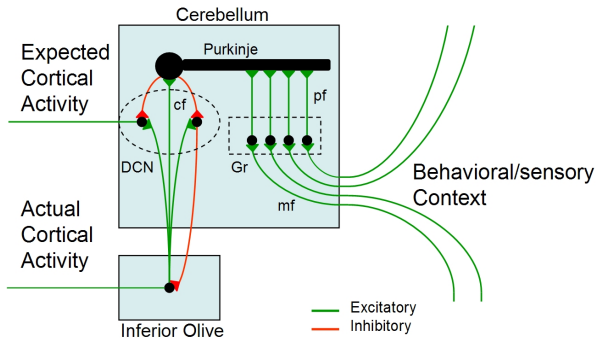


Fig. 2. Semantics of the inputs and output of the cerebellum.

from a neuron in the motor areas to the striatum, through the direct and indirect pathways of the basal ganglia to the ventral lateral nucleus of thalamus, and back to the original neuron. The other loop is from a neuron in the motor areas through the parvocellular neurons of the red nucleus to the inferior olive, through a climbing fiber, Purkinje cell and deep cerebellar nucleus neuron to the ventral lateral nucleus, and back to the original neuron. Both of these loops influence cortical activity through thalamic projections.

Third, the granule cells of the cerebellum receive massive afferents from the spinal cord, the cranial nerves, and broad areas of both the behavioral and sensory cortex² through the pontine nuclei.

3 Proposed Role for the Cerebellum

This paper proposes that the cerebellum’s role is to drive cortical activity at the statistically expected, nominally desired level for an organism’s given context. This context is quite rich and it includes sensory, vestibular, and proprioceptive information from the spinal cord and cranial nerves, and the cortical representation of the behavioral, phenotypic, and environmental state via the pontine nuclei. As shown in Fig. 2, the context is presented to the cerebellar granule cells via mossy fibers and to the Purkinje cells via parallel fibers. Fig. 2 also shows that the olivary-cerebellar subsystem is modulated by actual ongoing cortical activity. Section 5 describes how the cerebellum model uses context and cortical activity information to learn the proper output.

Since the cerebellum is a driver of cortical activity through the thalamus and adjusting its output would shift the expected activity, an apparent inconsistency exists. This inconsistency can be resolved by adjusting the contributions of the other drivers of cortical activity. These other drivers include cortico-cortical projec-

²The parietal, occipital, and temporal lobes.

tions and basal ganglia outputs conveyed by the thalamus. Since cortico-cortical learning is slow and cortico-cortical afferents partly define the context the cerebellum is mapping, changes in basal ganglia outputs are the primary means to accommodate cerebellar learning. The basal ganglia presumably accomplish this through changes in its cortical and limbic system afferents derived from an organism’s behavioral objectives and interaction with the environment.

In this model, therefore, the cerebellum is responsible for producing expected cortical activity for a given context and the basal ganglia are responsible for producing variations on expected behavior, including novel behaviors, refinements of movement, etc. As specific behavioral variations become “expected,” the cerebellum takes them over. Since both the cerebellum and basal ganglia project to the VLN and the VLN projects to the cortex, the cortex cannot distinguish the source of its afferent activity and is oblivious to this transfer of responsibility.

4 The Model Description and Dynamics

Excluding the basal ganglia, the scope of the model includes the entire circuit shown in Figure 1. The basal ganglia outputs, however, are modeled. The neurons are implemented using Wilson-Cowan type models [8]. If the firing rate of each neuron in the circuit is represented as a dimension of a vector, \mathbf{v} , then the neural dynamics are represented as:

$$\dot{\mathbf{v}} = \beta(W \cdot \mathbf{v} + \mathbf{v}_0 - \mathbf{v}), \quad (1)$$

where W represents the synaptic weight matrix, \mathbf{v}_0 represents the intrinsic firing rates of the neurons, and β is a constant controlling the rate of change. The intrinsic firing rates and the weight matrix are provided in Table 1. As indicated, only the inferior olive and deep cerebellar nucleus neurons have non-zero intrinsic firing rates and only the parallel fiber weights are plastic.

The output O is interpreted as the total spinal projection; this is the sum of cortical and red nucleus magnocellular activity.

To mimic the complex, interactive, corrective contributions of the basal ganglia to cortical activity, the basal ganglia output evolves as follows:

$$\dot{b} = \gamma(D - O), \quad (2)$$

where γ is a time constant, D is the damped desired output, $\dot{D} = \zeta(d - D)$, d is the desired output, and ζ is the damping rate constant. For convenience, the output of the basal ganglia is treated as excitatory.

The context values are represented as dimensions of \mathbf{v} and are produced by the experimental setup. The context values are damped like D to accommodate binary contexts.

Table 1. Intrinsic Firing Rates and Connection Weights

#	Source	v_0	Synaptic Weight with Target										
			1	2	3	4	5	6	7	8			
1	Cortex			1.0									
2	RN _p					1.0							
3	RN _m												
4	IO	0.3					0.3	1.0	1.0				
5	PK							-2.0	-2.0				
6	DCN _e	1.0			1.0								1.0
7	DCN _i	1.0				1.0							
8	TH		1.0										
9	BG												1.0
10	Context							*					

Empty cells have a value of 0

* Modified through learning

5 Cerebellar Learning

In this model, learning only takes place at the parallel fiber/Purkinje cell synapses. This learning must enable the cerebellum to drive expected cortical activity using only ongoing cortical activity as a training signal.

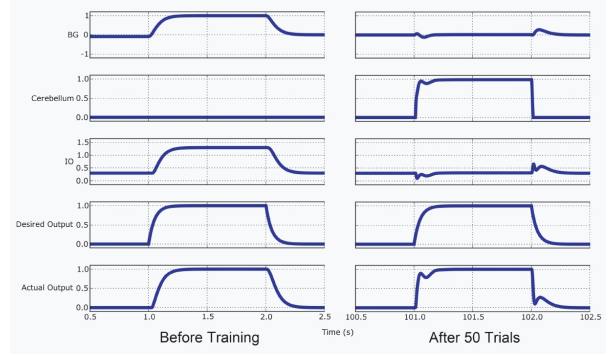
This is achieved using a cerebellar learning mechanism first described by Kenyon, et al. [5, 6]. This mechanism assumes that parallel fiber synapses undergo LTD when coactive with climbing fiber activity and LTP when active without coincident climbing fiber activity. Since Purkinje cells (PK) disinhibit inferior olive activity (IO), as Fig. 2 shows, a negative feedback loop is established. An increase in IO activity will cause LTD of parallel fiber synapses. This will result in less IO disinhibition due to lessened PK activity and IO activity will drop. If it drops too low, LTP will result in more disinhibition and IO activity will increase. This mechanism will maintain IO activity at a stable firing rate, observed to be 1–4 Hz, where LTP and LTD effects are balanced.

The learning rule for a Purkinje cell can be expressed

$$\dot{w}_j = \eta p_j (\alpha - \xi c), \quad (3)$$

where w_j is the weight of the j^{th} parallel fiber input, η is the learning rate, p_j is the fiber's activity level, α is the stable IO firing rate, c is the climbing fiber activity level, and ξ is a constant to balance c activity with α .

By coupling PK inhibition to the excitatory output of the deep cerebellar nuclei (DCN), a cortical tracking capability is established. Assume an initial equilibrium point where cortical inputs from the red nucleus and disinhibitory inputs from the PK drive the IO at a stable firing rate and the DCN outputs drive cortical activity at the expected level for a given situation. If other influences, such as the basal ganglia, cause an increase in IO activity, LTD will reduce DCN inhibition while driv-

**Fig. 3.** Associative learning.

ing the IO to its stable firing rate. This will cause DCN outputs to increase. Similarly, the DCN outputs will decrease with decreased IO activity. In both cases, a new equilibrium will be established with the cortex driven by the cerebellum at the expected activity level.

6 Associative Learning

One objective of the model was to explain how a behavioral response can be associated with a specific context, such as a specific stimulus. To test whether the model has this capability, two values were used to indicate two contextual states or contexts: 01 and 10. The system was constructed with a single cerebellar loop and initialized to produce a single output of 0 for both contexts. Then, for periods of one second each, the two contexts were alternately presented.

To learn the association, cortical activity was driven by the basal ganglia during context 10 to achieve a desired output of 1. For context 01, cortical drivers were withdrawn. As shown in Fig. 3, the basal ganglia initially drove cortical activity to produce the desired output. The cerebellum did not initially contribute at all. The IO activity, however, increased significantly, which drove the learning process. After 50 trials, however, the cerebellum learned to drive the proper cortical activity and the basal ganglia no longer contributed. Thus, the means to produce the desired output was effectively shifted from the basal ganglia to the cerebellum. A new association was established between context 10 and the desired output of 1. Finally, the IO activity level returned to the stable firing rate for both contexts after training.

7 Motor Learning

To test motor learning capabilities, Katayama's arm model with two joints and six muscles was used [4]. See Fig. 4. Six cerebellar loops were combined to control the muscles. The desired output was the combination of mo-

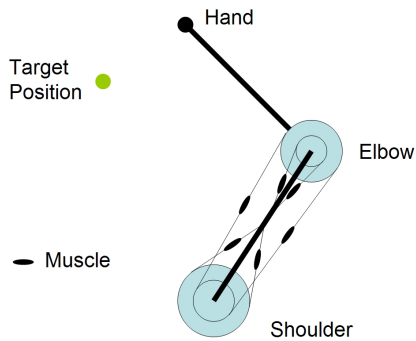


Fig. 4. Arm model with six muscles and a representative reaching target position.

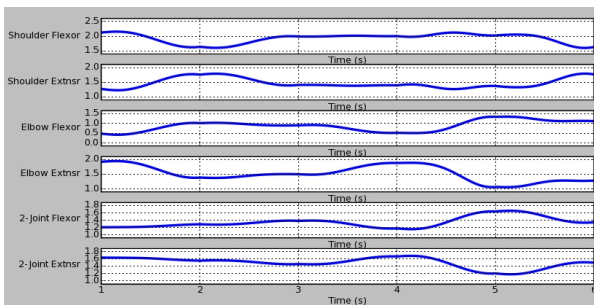


Fig. 5. Arm muscle commands over five successive reaches of one second duration.

tor commands necessary to reach each random targets in a 2D space with minimum jerk trajectories. To illustrate the challenge, the required muscle commands for five successive reaches are shown in Fig. 5. A six dimensional space was constructed using the joint angles, the joint velocities, and the target offset. This space was decomposed using radial basis functions to create the context representation. Finally, the cerebellum was trained for 80 one second reaches to random target positions.

Fig. 6 shows that the basal ganglia contribution to motor control diminishes approximately exponentially with time and that this is made up by the cerebellum. The bottom trace shows that the movement error does not get worse as the transfer takes place. After training, the cerebellum is almost entirely responsible for achieving minimum jerk trajectories using a multijointed, multi-musled arm.

8 Conclusions

A model of the cerebellum has been presented that is capable of associative learning and motor control. It is

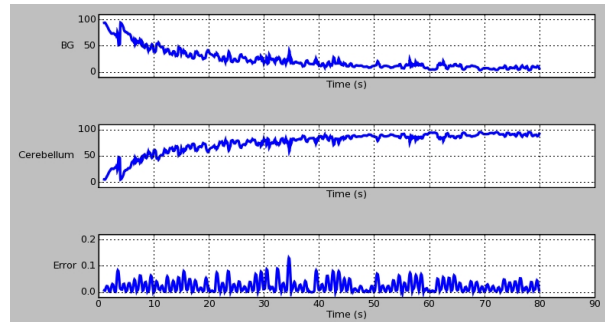


Fig. 6. Exponential transfer of cortical drive activity from basal ganglia to the cerebellum in learning of complex movements.

consistent with known biological constraints. The model depends upon and explains the cerebellum's relationships with other structures in the brain.

The model also supports a more general interpretation of the cerebellum's role. That is, the cerebellum's role is to drive the expected, nominally desired cortical activity for each given situation. Such a capability frees other brain structures and processes to respond to novel and refined aspects of behavior. It also enables capabilities that slower, more expensive processes could not achieve, such as ballistic movements.

Future work includes more realistically modeling the information provided by proprioception and the way it is presented to the cerebellum; and using recursive representations to achieve location independence of gesture execution and learning, and to investigate support for "cognitive" functions.

References

- [1] Afifi A.K., & Bergman R.A., (2005) *Functional Neuroanatomy*, 2nd Ed.. New York: McGraw-Hill
- [2] Hikosaka O., et al. (1998) *Neurobiology of Learning and Memory* 70, 127–149
- [3] Kandel E.R., et al. *Neural Science*, 4th Ed., New York: McGraw-Hill
- [4] Katayama M., & Kawato M., (1993) *Biological Cybernetics* 69, 353–362
- [5] Kenyon G.T., Medina J.F., & Mauk M.D., (1998) *Journal of Computational Neuroscience* 5, 17–33
- [6] Kenyon G.T., Medina J.F., & Mauk M.D., (1998) *Journal of Computational Neuroscience* 5, 71–90
- [7] Thompson R.F., (1988) *Trends in Neuroscience* Apr;11(4):152–5.
- [8] Wilson H.R. & Cowan J.D. (1972) *Biophysical Jnl*, 12:1–24.