

A robot conditioned reflex system modeled after the cerebellum

by JAMES S. ALBUS

*National Aeronautics and Space Administration
Greenbelt, Maryland*

INTRODUCTION

Most modern theories of behavior involve a hierarchical structure whereby low level behavioral units are controlled or manipulated by higher centers so as to produce characteristic patterns of movement. In the simplest life forms, low level behavioral units may consist of simple reflex arcs with very little higher level control. In intermediate forms, the low level behavioral units may be relatively complex in themselves and subject to sophisticated control from higher centers. In the most advanced nervous systems, higher centers may themselves be arrayed in a hierarchical structure, with each level monitoring activity and exerting control over the levels beneath it.

It has been suggested^{1,2,3} that the brain has a repertoire of behavioral units arranged much as the keys of a piano. Friedman suggests that a higher level "selecting mechanism activates these behavioral units to produce complex behavior just as an accomplished piano player produces a Beethoven sonata from his simple keyboard."

Theoretical workers in the behavioral sciences have suggested a number of ways in which these selecting mechanisms might interact to choose the proper behavioral units for the task to be accomplished. Just what these behavioral units are, however, or how they are specifically controlled, has been an open question.

There have been, of course, a number of hypotheses concerning the structure and function of reflex arcs. At a very simple level, such as the motor system "gamma loop," these models have been convincing. However, at more complex levels, theories such as Hebb's cell assemblies⁴ have been completely unsuccessful in providing substantive explanations for behavioral phenomenon. Sufficient quantitative data concerning the anatomy and physiology of complex brain structures has, until recently, simply not been available for formu-

lating precise models with convincing properties. In the absence of data, most models have been vacuous conjections.

In the past 8 to 10 years, however, the electron microscope and refined techniques of microneurophysiology have revealed quantitative data of considerable detail concerning the structural and functional organization of the brain, particularly in the cerebellum. A great deal of the physiological data about the cerebellum has come from an elegant series of experiments by Eccles and his co-workers. These data have been compiled along with the pertinent anatomical data, in book form by Eccles et al.⁵ This book set forth one of the first reasonably detailed theories on the function of the cerebellum.

Shortly after the publication of Eccles' book, another theory was developed by two different researchers working independently. Marr⁶ published his Theory of Cerebellar Cortex in 1969 and shortly thereafter the present author⁷ published a Theory of Cerebellar Function. Recently this theory has been developed further and reduced to computer software for the control of a mechanical manipulator.

FUNCTION OF THE CEREBELLUM

Although the Theory of Cerebellar Function was developed largely from neurophysiology and anatomical evidence, its reduction to computer software can be explained without detailed knowledge of the biological literature.

The cerebellum, along with the higher level brain centers which control it, can be thought of as a type of finite state machine.

$$M = (S, I, O, \delta, \lambda)$$

where

S is a finite non-empty set of states

I is a finite non-empty set of inputs
 O is a finite non-empty set of outputs
 $\delta: SXI \rightarrow S$ is the transition function
 $\lambda: SXI \rightarrow O$ is the output function

The "set" (or state) of the higher level brain centers determines the state S of the cerebellum. The sensory signals from various nerve endings in the limbs being controlled provide input I . The combination of I impinging on the cerebellum in state S , produces output O . The output function $\lambda: SXI \rightarrow O$ corresponds to a reflex arc. In the cerebellum the function λ is defined, and may be altered, through the process of learning.

The transition function $\delta: SXI \rightarrow S$, although undoubtedly of great importance to theories of higher level perception and intelligence, is considered beyond the scope of the elemental reflex level control functions being addressed in this paper. We are here considering merely how the cerebellum can be put into state S by the higher level centers, and then act as a reflex arc which transforms input I into output O under the operation λ . We will also discuss how λ can be altered through training.

Input I enters the cerebellum via mossy fibers from

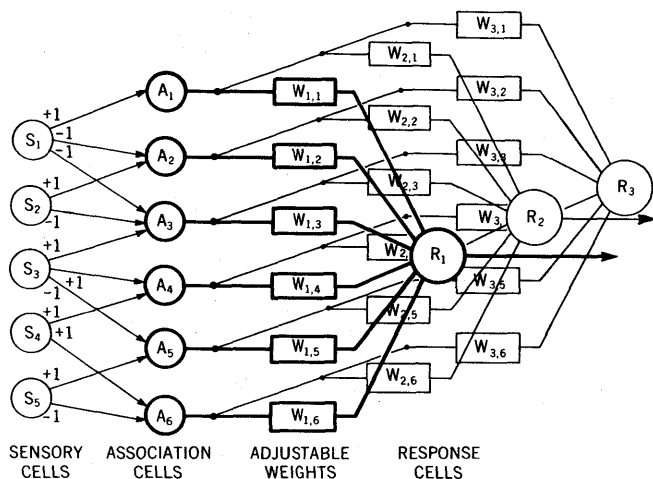


Figure 1—Classical Perceptron. Each sensory cell receives stimulus either +1 or 0. This excitation is passed on to the association cells with either a +1 or -1 multiplying factor. If the input to an association cell exceeds 0, the cell fires and outputs a 1; if not, it outputs 0. This association cell layer output is passed on to response cells through weights $W_{i,j}$ which can take any value, positive or negative. Each response cell sums its total input and if it exceeds a threshold, the response cell R_j fires, outputting a 1; if not, it outputs 0. Sensory input patterns are in class 1 for response cell R_j if they cause the response cell to fire, in class 0 if they do not. By suitable adjustment of the weights $W_{i,j}$ various classifications can be made on a set of input patterns.

(Figures 1, 2, and 3 reprinted by permission from Mathematical Biosciences 10, 1971)

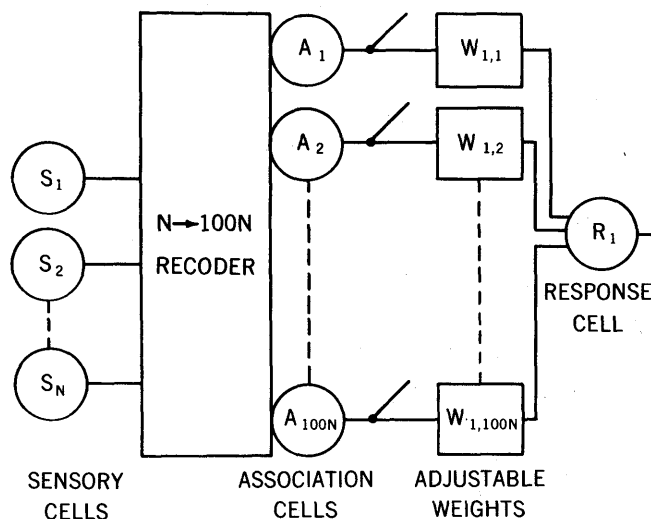


Figure 2— N , $100N$ Expansion Recoder Perceptron. The association cell firing is restricted such that only 1-2 percent of the association cells are allowed to fire for any input pattern. This Perceptron has a large capacity and fast learning rate, yet it maintains the number of association cells within limits reasonable for the nervous system

the periphery. (The engineer unfamiliar with anatomical nomenclature must excuse the quaint terminology. Many terms, like mossy fiber, were coined by early anatomists over a century ago. Peering through their crude microscopes and seeing fibers resembling moss, they merely called them as they saw them. In other instances, features such as Purkinje cells were named after the first investigator who observed them.) All mossy fiber input enters a section of the cerebellum called the granular layer. In the granular layer, information carried by the mossy fibers in the form of pulse interval (or frequency) modulation is transformed into information carried on parallel fibers. The important feature of this transformation is that there are from 100 to 1000 times as many parallel fibers coming out of the granular layer as there are mossy fibers going in. This implies that in the granular layer, information is recoded. The evidence seems to indicate that the granular layer transforms mossy fiber information in the frequency domain into parallel fiber information in the spatial domain. The theory predicts that only a very few (i.e., about 1-2 percent) parallel fibers are active for any given pattern of pulse frequency modulation on mossy fibers.

Output from the cerebellum itself is via Purkinje cells. The theory predicts that Purkinje cells perform a weighted summation of parallel fiber activity analogous to the way in which a Perceptron response cell performs a weighted summation on association cell firings. Thus

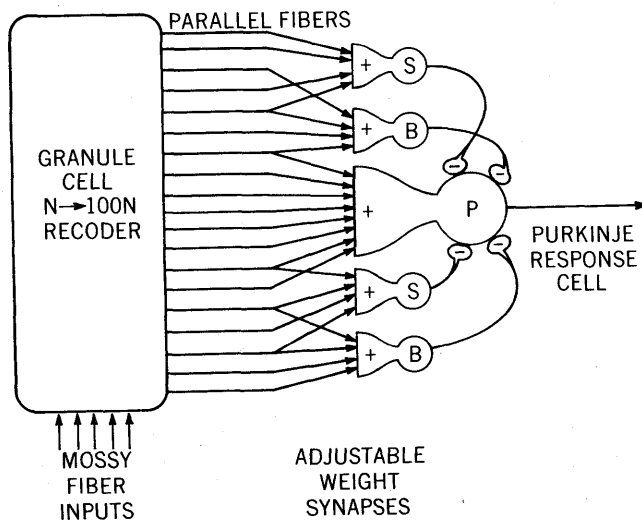


Figure 3—Cerebellar Perceptron: *P*, Purkinje cell; *B*, basket cells; *S*, stellate *b* cells. Each Purkinje cell has inputs of the type shown

the cerebellum can be considered to be a form of Perceptron where mossy fiber input is analogous to sensory cell firings. The granular layer corresponds to the interconnection network between sensory cells and association cells, parallel fibers correspond to association cell outputs and the synaptic connections between parallel fibers and Purkinje cells correspond to the variable weights. The Purkinje cells themselves correspond to the Perceptron response cells. This analogy can be seen in Figures 1, 2, and 3. Figure 1 is the classical Perceptron. Figure 2 shows the classical Perceptron modified to conform to the anatomical fact that the cerebellar granular layer contains more than 100 times as many parallel fibers as input mossy fibers. Figure 3 extends the Perceptron analogy to take into consideration the fact that in the nervous system certain types of cells are excitatory and other types are inhibitory, but no type is both. Thus, in order for the cerebellar Perceptron to have both positive and negative valued weights connecting parallel fibers to Purkinje cells, some intermediary cell types (i.e., *B*, basket cells, and *S*, stellate *b* cells) are necessary. To an engineer, these intermediary cells are inverters.

There is one important difference between the classical Perceptron and its cerebellar counterpart. The classical Perceptron typically accepts only binary input signals, performs an analog weighted summation, compares this sum with a threshold, and responds with a binary output. The cerebellar Perceptron, on the other hand, accepts input which, although consisting of binary pulses, contains information in the form of pulse frequency modulation which is essentially analog

in nature. This analog data is recoded from the frequency domain to the spatial domain by the granular layer. The Purkinje response cell, at least to a first approximation, can be considered a linear summation device. It performs no thresholding in the sense of a classical Perceptron response cell. Purkinje cells are typically spontaneously active at some steady-state output rate. A weighted summation of parallel fiber activity merely increases or decreases the frequency of the Purkinje output pulse train. Thus, both outputs and inputs to the cerebellum should be considered to be analog signals coded into pulse frequency modulation.

Learning

The cerebellum is hypothesized to learn by an error correction system similar to Perceptron training algorithms. Each Purkinje cell is contacted by a single climbing fiber. These climbing fibers are hypothesized to carry the information necessary to adjust synaptic weights in an error correcting manner. Climbing fibers carry information from higher motor centers as well as centers of emotional reward and punishment.² These higher centers presumably are able to sense conscious motor commands, compare these conscious commands against the cerebellar reflex motor output, and correct the cerebellar output when it deviates from what the higher centers consider to be satisfactory performance. This correction takes place by adjusting the synaptic weights between active parallel fibers and erroneously responding Purkinje cells. The weights are adjusted so as to null the difference between what conscious centers send to the motor system, and what the cerebellar reflex produces. Thus, as training proceeds, more and more of the routine motor control can be relegated to the cerebellar reflex arc, and higher centers are then free to concentrate on other matters.

This corresponds to the common experience which everyone has had when learning a new motor skill. At first, a task such as driving an auto, playing a musical instrument, or roller skating requires a great deal of conscious concentration. However, as learning proceeds, more and more of the new motor skill comes under reflex control, and less conscious mental effort is required. This presumably is the process of training the cerebellum (and other similar subconscious motor centers) to take over the repetitive and routine tasks which can be controlled by reflex responses.

The cerebellum thus can also be viewed as a memory. The mossy fibers constitute the address lines. The climbing fibers constitute the data storage inputs. And the Purkinje cell outputs correspond to the contents of the memory. This is illustrated in Figure 4. The mossy

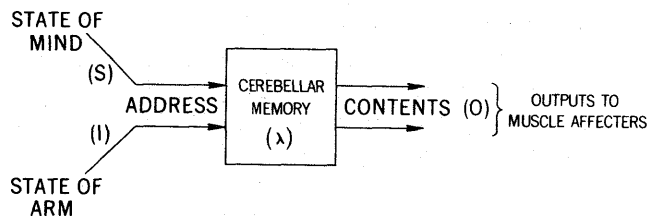


Figure 4—Cerebellar Memory. The state S of higher brain centers is communicated along with the input I from peripheral proprioceptors to the cerebellum via mossy fibers. S and I together constitute an address. Data to be stored arrive via climbing fibers

fiber input constitutes an address. The Purkinje cell response corresponds to the memory contents. Each Purkinje cell output can thus be considered a separate memory bank. By this means the cerebellum achieves the redundancy which is so characteristic of circuitry in the brain.

It, of course, is obvious that if mossy fiber address lines are essentially analog in nature, there exists an enormous number of possible addresses. If we assume that there are N mossy fiber address lines, and each mossy fiber can carry an analog signal with 50 distinguishable values of pulse frequency, then we have 50^N possible addresses. If we consider that each square millimeter of granular layer has approximately 5×10^4 mossy fibers entering it, we clearly have a potentially enormous number of addresses. However, one must remember that if the world is subjected to a state-space analysis, there exists an equally enormous number of possible states-of-the-world. Mossy fiber input from sensory receptors in the limbs are essentially reporting the state of the limbs. Since people and animals are able to cope with the infinity of possible states in the real world, it is clear that somehow these states are grouped into a manageable number of sets of states. States within such groupings are for all practical purposes equivalent. So too, the virtual infinity of possible mossy fiber addresses are grouped into sets of essentially equivalent addresses. This grouping is accomplished by the granular layer. The granular layer performs a transformation such that if two mossy fiber addresses are within an equivalence group, the same pattern of parallel fiber outputs will occur.

The mossy fiber input can be considered a vector

$$I = (mf_1, mf_2, mf_3, \dots, mf_N)$$

where mf_i is the firing rate of the i -th mossy fiber.

We can define similarity between mossy fiber patterns

in terms of the Hamming distance H_I between input vectors I and I' .

$$H_I = \sum_{i=1}^N |mf_i - mf'_i|$$

The mossy fiber input vector I is transformed by the granular layer into a parallel fiber vector

$$J = (pf_1, pf_2, pf_3, \dots, pf_{100N})$$

where pf_i is the firing rate of the i -th parallel fiber.

The theory hypothesizes that at any instant of time only about two percent of the pf_i firing rates are non-zero. Thus, the vector J is a very sparse vector. The principal feature of this transformation is the conversion of mossy fiber patterns in the frequency domain to parallel fiber patterns in the spatial domain. Parallel fibers thus are hypothesized to code information in terms of the specific set of parallel fibers which have non-zero firing rates. We can define a set

$$L = \{pf_i \mid pf_i \text{ has a non-zero firing rate}\}$$

We can then represent similarity between two parallel fiber patterns J and J' in terms of the intersection

$$L \cap L'$$

The granular layer is hypothesized to perform such that if H_I is small, $|L \cap L'|$ will be large, and as H_I grows large $|L \cap L'|$ will approach zero. This implies that training for dissimilar tasks (i.e., such that H_I is large) will produce very little interference. Weights adjusted for pattern I will be entirely different from those adjusted for pattern I' because $|L \cap L'|$ is zero. However, for similar patterns, training will generalize. Similar mossy fiber patterns (i.e., H_I small) will cause many or most of the same weights to be adjusted because $|L \cap L'|$ is large. Thus, the cerebellum need not be trained to cope with every possible mossy fiber address corresponding to every possible state of the arm. Instead, training over a small but representative sample of the possible states will suffice.

An electro-mechanical model

Consider now how such a model of the cerebellum can be reduced to computer software. An IBM 1800 computer was connected to a Rancho Los Amigos arm with seven degrees of freedom. Each degree of freedom was driven by a separate motor. Each motor amplifier was controlled by a computer model of a single Purkinje summation cell. Each Purkinje summation in the computer thus represents a large number of Purkinje cells in the real cerebellum; some activating flexor muscles, others activating extensor muscles.

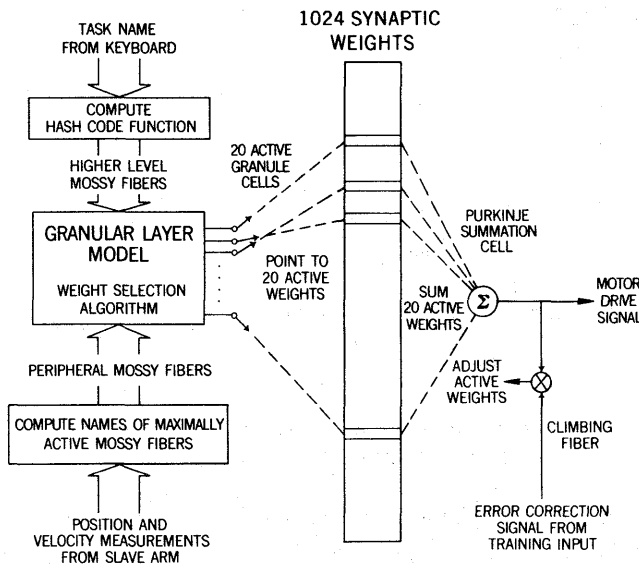


Figure 5—Computer model of cerebellar Perceptron

Each of the seven Purkinje summation cells in the computer are related to a table of 1024 synaptic weights, as shown in Figure 5. These weights are adjustable over the range -32767 to $+32767$. According to the theory, only two percent of the parallel fibers, and hence two percent of the synaptic weights, are activated at any one time. In the computer model, the granular layer selects 20 out of the 1024 weights to be active. The Purkinje cell then sums the values of these 20 active weights. This summation is the Purkinje cell output.

In the model, mossy fiber inputs convey information concerning the position and velocity of each joint. This information constitutes an address which is converted by the granular layer into a set of 20 active parallel fibers. These 20 parallel fibers connect to 20 active weights which are summed by the Purkinje cell. The value of this summation can be considered to be the "contents" of the memory location addressed by the mossy fiber pattern.

For any position-velocity state of the arm, each Purkinje summation cell delivers a drive voltage to the actuator motors. If this voltage is not appropriate to the task being attempted, it can be modified by adjusting the 20 active weights in each Purkinje summation. This is the training mode. When the arm is being trained to perform a particular task, the weights selected by the granular layer are adjusted by the training algorithm to follow the instructions being generated by the teacher. The teacher in the model is a master arm worn by a human operator. The training operation begins by the operator entering the name of the task

to be learned on a keyboard. This name corresponds to the psychological "set" or state of the higher centers in the brain. This determines the state S of the cerebellum. For example, the number 0101 on the keyboard might correspond to the task "reach-out." The operator would then proceed to teach the cerebellar model by performing a reach-out motion with the master arm. At closely spaced intervals along the reach-out trajectory the controlled arm position is compared against the master arm position. Whenever a discrepancy is detected, the weights connected to the 20 active parallel fibers are adjusted so as to drive the motors in a direction which will null the difference.

By this means, the memory stores the proper motor drive voltage for each position-velocity state along the desired trajectory. Repeated training can store the proper corrective voltage outputs for other states to either side of the desired trajectory. The generalization properties of the memory make it feasible to train the arm on only a representative sample of the universe of possible states, and still achieve satisfactory performance. The process of training defines the function λ for the universe of input states I encountered in performing the task S .

MODELING THE GRANULAR LAYER

The selection of which set of parallel fibers are active at any instant of time is the function of the granular layer. It is one of the principal hypotheses of the Theory of Cerebellar Function that the manner in which this selection is made gives the cerebellum its unique powers of motor coordination, precision control, and flexibility.

The origin of coordination

It has been experimentally shown⁸ that a somatotopic mapping exists from the cerebellar cortex to the muscles of the body. This means, for example, that Purkinje cells affecting the elbow are likely to be physically located in close proximity to each other, and an appreciable distance from those affecting the wrist. Since any single parallel fiber extends only about 1 mm. along a folial ridge, it is quite unlikely that a parallel fiber which contacts an elbow Purkinje will also contact a wrist Purkinje. This implies that the sets of parallel fibers involved in Purkinje summations for controlling specific joints in the model should be disjoint. Thus, in the model, each Purkinje summation will involve a separate set of granule cells and a separate set of weights.

On the other hand, somatopy is much less well defined from the periphery to the cerebellum.^{9,10,11,12}

Granule cells which are close neighbors in the cerebellum often have receptive fields at widely separate areas of the same limb or even in different limbs.¹³ Mossy fibers enter the cerebellum and ramify diffusely throughout a single folia and even into several different folia. This is not to say that somatopy is non-existent for mossy fiber input, just that it is diffuse and overlapping.

The implication is that Purkinje cells are fairly specific in their control over individual muscles, or synergistic groups of muscles. However, the input to any particular Purkinje cell in the cerebellum, while strongest from its somatopic area in the periphery, is also appreciably strong from other areas of the periphery. Strength of influence from neighboring peripheral areas falls off slowly with distance. Thus, the strongest input to a Purkinje cell controlling the elbow should arrive via mossy fibers from the elbow. However, an appreciable input to the elbow Purkinje should also come from the forearm and shoulder, and to a lesser extent from the wrist and hand. Similar conditions exist for each set of Purkinje cells corresponding to each joint. Input should be strongest from the joint to which a Purkinje projects, and fall off in strength from other joints as a function of distance.

The relevance matrix

In order to model the relative degree of influence which mossy fibers from the various joints have on the sets of granule cells unique to each joint, a relevance matrix is constructed as shown in Figure 6.

The numbers in this matrix indicate relative values. Each row sums to 72. (The number 72 derives from the fact that there are 72 entries in the matrix shown in Figure 8). The first row of the matrix suggests that 30/72 of the mossy films influencing shoulder rotation carry feedback information concerning the state of the shoulder rotation joint, 12/72 carry information con-

Input to Purkinje controlling	via Mossy fiber from					
	Shoulder Rotate	Shoulder Lift	Elbow Rotate	Elbow Lift	Forearm Rotate	Wrist Lift
Shoulder Rotate	30	12	9	9	6	6
Shoulder Lift	12	30	9	9	6	6
Elbow Rotate	6	6	30	15	9	6
Elbow Lift	6	6	15	30	9	6
Forearm Rotate		6	9	9	30	12
Wrist Lift		6	6	6	12	30
Finger Grasp		6	6	6	9	12

Figure 6—Relevance Matrix. This matrix represents the relative degree to which input from each joint is relevant to the computation of motor output for each joint

FIRING RATE

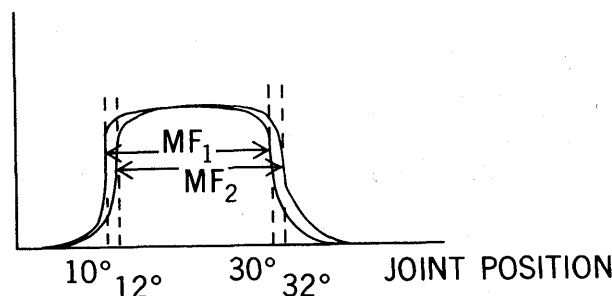


Figure 7—Assumed mossy fiber firing rates plotted against joint position for two different mossy fibers

cerning the shoulder lift joint, 9/72 concern the elbow rotation joint, 9/72 concern the elbow lift joint, 6/72 concern forearm rotation, and 6/72 concern wrist lift. Similarly for forearm rotation. Row 5 of the matrix indicates that 30/72 of the peripheral mossy fiber input carries information about forearm rotation, 12/72 about wrist lift, 9/72 about elbow lift, 9/72 about elbow rotation, 6/72 about shoulder lift, and 6/72 about finger grasp. Each functional portion of the cerebellum has a different mixture of inputs. In each case 30/72 of the input to the control circuit for each joint is simply feedback information from that joint. The remaining 42/72 of the input carries information concerning related joints.

The mechanism of selection

It seems to be the case¹⁰ that individual mossy fibers fire at their maximal rate when specific conditions exist in specific parts of the periphery. A mossy fiber carrying joint position information will tend to fire at its maximum rate when a specific joint is within a certain range of positions. For example, a typical elbow position fiber might fire at its maximum rate when the elbow joint angle is between 10° and 30°, and at a slower rate otherwise. A different position fiber might fire maximally for elbow positions between 12° and 32°, etc. It has been observed that position fibers fire maximally over some extended range and that considerable overlap exists between the maximal firing ranges of various fibers. See Figure 7.

A mossy fiber carrying information concerning joint velocities will tend to fire at its maximum rate when a particular joint is moving at a rate within a certain range of velocities. Some mossy fibers indicate positive velocities and others negative velocities.

Mossy fibers which fire at their maximum rate are of

critical importance if it is true that only 1-2 percent of the parallel fibers are active at once. This 1-2 percent hypothesis implies that for any granule cell to be active a very special set of excitation conditions must be satisfied. Active granule cells must have their input in the upper 1-2 percent of excitation values. Since granule cells have relatively few mossy fiber inputs, each input contributes a large percentage to the total excitation of the cell. It is thus reasonable to assume that for a granule cell to become a member of the very select set of active cells, all, or nearly all, of its mossy fiber inputs must be firing at or near their maximum rate.

From anatomical measurements concerning the numbers and densities of granule cell inputs,⁵ and arguments concerning probability of excitation by mossy fiber inputs,¹⁴ it is possible to predict that:

- 10 percent of active granule cells have 1 input (*s*)
- 20 percent of active granule cells have 2 input (*s*)
- 20 percent of active granule cells have 3 input (*s*)
- 20 percent of active granule cells have 4 input (*s*)
- 15 percent of active granule cells have 5 input (*s*)
- 10 percent of active granule cells have 6 input (*s*)
- 5 percent of active granule cells have 7 input (*s*)

This implies that in a model where 20 granule cells are active,

- 2 should be a function of 1 mossy fiber (*s*)
- 4 should be a function of 2 mossy fiber (*s*)
- 4 should be a function of 3 mossy fiber (*s*)
- 4 should be a function of 4 mossy fiber (*s*)
- 3 should be a function of 5 mossy fiber (*s*)
- 2 should be a function of 6 mossy fiber (*s*)
- 1 should be a function of 7 mossy fiber (*s*)

Notation for naming

In order to compute the seven sets of 20 active granule cells in a computationally efficient manner, it is convenient to introduce some special notation. First, each mossy fiber entering the cerebellum will be given a unique number. Such a numbering is, in fact, a notation for naming. We will refer to each mossy fiber's number as its name. Thus, mossy fiber #1 is named 1, mossy fiber #2 is named 2, etc. It will also be convenient from time to time to refer to mossy fibers by another convention, or "nickname." 1 will be nicknamed MF_1 , 2 will be nicknamed MF_2 , etc.

We will now define a classification of mossy fibers called an exclusive set.

Df: An exclusive set is the set of all mossy fibers such that no two mossy fibers can possibly be maximally active simultaneously.

For example, if

MF_1 is maximally active when the elbow is between 0° and 40°

MF_2 is maximally active when the elbow is between 40° and 80°

MF_3 is maximally active when the elbow is between 80° and 120°

MF_4 is maximally active when the elbow is between 120° and 160°

then $\{MF_1, MF_2, MF_3, MF_4\}$ is an exclusive set.

Df: A complete exclusive set is an exclusive set in which at least one mossy fiber is always maximally active. For example, the exclusive set given above would be a complete exclusive set if the elbow would never move outside the range 0° to 160° .

It is assumed that each joint has a number of both position and velocity mossy fibers with overlapping ranges of maximal excitation. These can be grouped into complete exclusive sets.

Df: ${}^iP_k^j$ is defined as the k th complete exclusive set of position-indicating mossy fibers coming from joint i and carrying information for Purkinje cell j .

Df: ${}^iV_k^j$ is the k th complete exclusive set of velocity mossy fibers from joint i going to Purkinje cell j .

Df: ${}^iP_k^j$ is the name of the mossy fiber in ${}^iP_k^j$ which is maximally active.

Df: ${}^iV_k^j$ is the name of the mossy fiber in ${}^iV_k^j$ which is maximally active.

For example, if

$${}^1P_3^2 = \{MF_{10}, MF_{12}, MF_{15}\}$$

and MF_{15} is maximally active, then

$${}^1P_3^2 = 15$$

The k subscript indicates different exclusive sets of mossy fibers with overlapping ranges. For example,

${}^1P_1^1$ might refer to the set $\{MF_1, MF_2, MF_3, MF_4\}$ such that

MF_1 is maximally active when $0^\circ \leq \alpha < 40^\circ$

MF_2 is maximally active when $40^\circ \leq \alpha < 80^\circ$

MF_3 is maximally active when $80^\circ \leq \alpha < 120^\circ$

MF_4 is maximally active when $120^\circ \leq \alpha < 160^\circ$

and

${}^1P_2^1$ might refer to the set $\{MF_5, MF_6, MF_7, MF_8\}$

such that

- MF_5 is maximally active when $0^\circ \leq \alpha < 38^\circ$
- MF_6 is maximally active when $38^\circ \leq \alpha < 78^\circ$
- MF_7 is maximally active when $78^\circ \leq \alpha < 118^\circ$
- MF_8 is maximally active when $118^\circ \leq \alpha < 158^\circ$

The granular layer matrices

The above notation will now make it possible to compute which 20 granule cells are selected for each Purkinje summation.

As was previously discussed, the statistical distribution of the number of inputs per active granule cell indicates that out of 20 active granule cells, two depend on only one mossy fiber input, four depend on two inputs, etc. This functional relationship can be formulated into a matrix as shown in Figure 8.

This matrix is the granular layer matrix for the shoulder rotation joint. Each space in the matrix is assigned to a complete exclusive set of mossy fibers such that the corresponding matrix element is the name of the maximally active mossy fiber in that set. All of the mossy fiber sets represented in this matrix are carrying information to the shoulder rotation Purkinje summation. In the model there are six other matrices, similar to this one, for the six other Purkinje summations. These matrices contain information concerning the state of the arm as reported by the peripheral mossy fibers.

The numerical values of the elements in these matrices change as the state of the arm changes. A small change in the arm will cause a small number of elements to change in value. A large change in the

⁶ V ₄
⁶ P ₄ ⁶ P ₈ ⁶ P ₁₂
⁵ P ₄ ⁵ P ₈ ⁵ P ₁₂ ⁵ P ₁₆ ⁵ V ₄ ⁵ V ₈
⁴ P ₄ ⁴ P ₈ ⁴ P ₁₂ ⁴ P ₁₆ ⁴ P ₂₀ ⁴ V ₄ ⁴ V ₈ ⁴ V ₁₂ ⁴ V ₁₆
³ V ₂ ³ V ₆ ³ V ₁₀ ³ V ₁₄ ³ V ₁₈ ³ P ₂ ³ P ₆ ³ P ₁₀ ³ P ₁₄ ³ P ₁₈
² V ₂ ² V ₆ ² V ₁₀ ² V ₁₄ ² V ₁₈ ² V ₂₂ ² P ₂ ² P ₆ ² P ₁₀ ² P ₁₄ ² P ₁₈ ² P ₂₂
¹ V ₂ ¹ V ₆ ¹ V ₁₀ ¹ V ₁₄ ¹ V ₁₈ ¹ V ₂₂ ¹ V ₂₆ ¹ V ₃₀ ¹ V ₃₄ ¹ V ₃₈ ¹ V ₄₂ ¹ V ₄₆
¹ P ₁ ¹ P ₂ ¹ P ₃ ¹ P ₄ ¹ P ₅ ¹ P ₆ ¹ P ₇ ¹ P ₈ ¹ P ₉ ¹ P ₁₀ ¹ P ₁₁ ¹ P ₁₂ ¹ P ₁₃ ¹ P ₁₄ ¹ P ₁₅ ¹ P ₁₆ ¹ P ₁₇ ¹ P ₁₈ ¹ P ₁₉ ¹ P ₂₀

Figure 8—Shoulder Rotation Granular Layer Matrix. This matrix is used to compute which set of 20 granule cells are active for the shoulder rotation joint. Note that input from shoulder rotation position ¹P_k and shoulder rotation velocity ¹V_k occupy a dominant role in the shoulder rotation matrix. In the model there are six additional granular matrices, one for each of the six other joints

state of the arm will cause many or all of the elements to change in value.

The particular assignments of elements from each joint represented in the matrices are derived, in part, from the relevance matrix in Figure 6. For example, the granular layer matrix for shoulder rotation, shown in Figure 8, has twenty position sets and ten velocity sets making a total of thirty sets of mossy fibers from the shoulder rotation matrix. This corresponds to the fact that the relevance matrix in Figure 6 specifies that 30/72 of the inputs to the shoulder rotation Purkinje should come from shoulder rotation mossy fibers. It is arbitrarily assumed that approximately $\frac{2}{3}$ of the inputs from each joint should be position indicators and the remaining $\frac{1}{3}$ should indicate velocity.

As can be seen from Figure 8, the assignment of particular sets to particular matrix elements was not done randomly. This was because it was felt that the number of matrix elements was too small to rely on statistical probabilities to give representative importance to the various mossy fiber inputs. Therefore, the various matrices were set up by hand and represent (as does the relevance matrix) the subjective judgment of the author as to which inputs are important to each Purkinje cell for controlling motor outputs. It is important to emphasize, however, that once these matrices are set up they are not changed. This corresponds to a granular layer structure defined by genetically coded interconnections and not structurally altered during an animal's lifetime.

Computation of active granule cell names

These granular layer matrices can now be used to compute which granule cells are active. In each matrix the 20 columns correspond to 20 active granule cells. Columns 19 and 20 correspond to granule cells with only one mossy fiber input. Columns 15, 16, 17, and 18 correspond to granule cells with two inputs, etc. The names of the active granule cells can be computed by the concatenation of elements in the columns of the matrices. For example, in Figure 8, the name of the

granule cell computed by column 11 would be ³P₉¹_n

²P₂¹ ¹P₁₁¹ If

$$\boxed{{}^3P_9^1} = 15, \boxed{{}^2P_2^1} = 12, \text{ and } \boxed{{}^1P_{11}^1} = 59$$

then, 151259 is the name of the granule cell computed by column 11. Thus, we have described a method for finding the names of granule cells, given the state of the mossy fiber inputs.

At this point in the discussion, the names we have

defined for granule cells are not yet in a particularly useful form nor are they necessarily even unique. The uniqueness problem is rather easily solved by requiring each mossy fiber name to contain the same number of digits. Leading zeros may be employed to accomplish this for mossy fiber names with small numerical values. For example, if it requires three digits to name all the mossy fibers, then mf_1 will be named 001, mf_2 will be named 002, etc. The question of how to utilize the names of granule cells once they are determined is slightly more complicated. Assume, for example, that an active granule cell is named 10956321. How do we find the contribution this cell firing makes to its respective Purkinje output? In the computer it is necessary to locate the weight which connects a granule cell to its respective Purkinje in order to compute its effect. This implies that for each granule cell which is active there must be a pointer which can locate its respective weight in a table of weights. Setting up a table of pointers for each possible granule cell name would be a most tedious job. Fortunately, there is a much simpler technique available. We can instead map the active granule cell names onto the set of integers from 1 to 1024 by means of hash-coding. This may be done quite simply by use of a pseudo-random number generator which uses the numerical value of the active granule cell name as an argument and computes a pseudo-random number in the range 1 to 1024. This pseudo-random number can be considered a new name or "alias" for the active granule cell. This alias can be used directly as a pointer to a table of 1024 weights which connect granule cells to Purkinje cells.

In summary, the following procedure obtains. Each joint has a matrix representing its own peculiar mossy fiber input distribution. In each of the seven matrices we may compute the mossy fibers which are maximally active from measurements of joint position and velocity. The names of these maximally active mossy fibers make up the elements in each matrix. Concatenation of the elements in each column yields the names of active granule cells. Each of these names are used as input to a pseudo-random number generator which maps them onto the integers from 1 to 1024. The result of this procedure is seven sets of 20 integers. These integers point to the 7 sets of 20 weights which are summed by the 7 Purkinje cells. The resulting summations define the output signals which drive the motors for each joint.

Computations for the mechanical arm

In the actual electro-mechanical arm there are, of course, no mossy fibers with overlapping characteris-

tics as in the physiological arm. Instead, each joint has a potentiometer which measures position to a rather high degree of precision. Such a measurement certainly contains all the information which a multiplicity of overlapping mossy fibers would contain. However, the system of overlapping mossy fibers and granule cell-Golgi cell network, produce the phenomenon that if the arm moves slightly, only a few granule cells change from active to inactive, or vice versa; the great majority of granule cells are unaffected. Mossy fibers map their activity into patterns of granule cell activity which are "nearly the same" when the state of the arm is "nearly-the-same." In order for the computer model to capture this "nearly-the-same" property, a method has been devised for converting the potentiometer readings into names of maximally active mossy fibers. These names can then be used as elements in the granular layer matrices.

Rather than attempt to describe the details of these computations in the limited space available here, the interested reader is referred to Reference 14 in the bibliography.

Input from higher centers

Higher level mossy fibers constitute a major source of input to the entire cerebellum. Once these fibers enter the granular layer, they are physically indistinguishable from peripheral mossy fibers. Because of their great numbers, they undoubtedly have a very strong influence on the selection of which granule cells are to be active. This would seem to imply that higher level mossy fibers should be represented in the matrices used to compute the granular layer transfer function. In fact, since the higher level mossy fiber input is so massive, it would seem that it should dominate the granular layer matrices. Surely a change in input on a mossy fiber system which so permeates the entire granular layer should affect, either directly or indirectly, the firing threshold of practically every granule cell in the cerebellum. And so it does. However, there is a simpler way of modeling the influence of higher level mossy fibers than inserting them in the granular layer matrices explicitly. A change in the pseudo-random number generator can model the effects of a very broad and diffuse change in granule cell thresholds throughout the entire cerebellum. Thus, the effect of higher level mossy fiber input can be modeled by assuming the hash-code operation to be under the control of higher centers.

The cerebral cortex is, of course, the place where decisions are made as to what task should be performed by the motor system. The cortex may decide that the

arm should perform the task "reach out." This "reach out" task would then be sent to the cerebellum as a specific cerebral mossy fiber firing pattern \bar{M}_{RO} . With the \bar{M}_{RO} pattern on the cortical mossy fibers and with the peripheral mossy fibers reporting the state of the various joints in the arm, the cerebellar Purkinje cells would tend to produce outputs to drive the motors to "reach out." If these outputs were incorrect, error correction signals via the climbing fibers would cause adjustments in the weights leading from the active granule cells. Thus, the cerebellum would be trained by error correction to correctly perform the task "reach out."

If the cerebral cortex were then to decide that the arm should perform the task "pull back," a new pattern \bar{M}_{PB} would be sent to the cerebellum via the cortical mossy fibers. This new pattern \bar{M}_{PB} would change the hash-code function and cause a completely different set of granule cells to be chosen for any pattern of peripheral mossy fiber inputs. Once again the cerebellum could be taught to perform the "pull back" operation by adjusting weights under direction of climbing fiber error correction. Each different task decided upon by the cerebral cortex can be communicated to the cerebellum by a different firing pattern on the cortical mossy fibers. In the model this implies that each different task should be assigned a different hash-coding function.

By this means the cerebral cortex is able to impose high level control on the motor system without worrying about continuous control of each individual muscle. The lower level control functions are carried out by the cerebellum. Only in the case of learning a new motor task, or in case of errors or deviations from the desired performance of previously learned tasks, does the cerebrum need to worry about the detailed control of lower level motor functions.

The cerebellum thus is the repository of detailed motor control sequences which have been previously learned under conscious effort, and which can be called up repeatedly by task name via higher level mossy fiber patterns much as subroutines are called by an executive program. The higher level mossy fibers make it possible for higher centers to control the cerebellar motor system with a macro command language.

REFERENCES

- 1 K Z LORENZ
The comparative method in studying innate behavior patterns
Physiological Mechanisms in Animal Behavior Symp Soc
Exper Biol 4 London Academic Press 1950 pp 221-268
- 2 J M R DELGADO
Free behavior and brain simulation
International Review of Neurobiology 6 C C Pfeiffer and
J R Smythies (Eds) New York Academic Press 1964
pp 340-449
- 3 L FRIEDMAN
Instinctive behavior and its computer synthesis
Behavioral Science 12 1967 pp 85-108
- 4 D O HEBB
The organization of behavior: A neuropsychological theory
New York Wiley 1949
- 5 J C ECCLES M ITO J SZENTAGOTHAÏ
The cerebellum as a neuronal machine
New York Springer-Verlag 1967
- 6 D MARR
A theory of cerebellar cortex
J Physiol London 202 1969 pp 437-470
- 7 J S ALBUS
A theory of cerebellar function
Mathematical Biosciences 10 1971 pp 25-61
- 8 A BRODAL
*Anatomical studies of cerebellar fiber connections with special
reference to problems of functional localization*
Progress in Brain Research The Cerebellum Vol 25 C A Fox
and R S Snider (Eds) Elsevier New York 1967 pp 135-173
- 9 B HOLMQUIST O OSCARSSON I ROSEN
Functional organization of the cuneo-cerebellar tract in the cat
Acta Physiol Scand 58 1963 pp 216-235
- 10 J K S JANSEN K NICOLAYSEN T RUDJORD
*Discharge pattern of neurons of the dorsal spinocerebellar
tract activated by static extension of primary endings of
muscle spindles*
J Neurophysiol 29 1966 pp 1061-1086
- 11 A LUNDBERG O OSCARSSON
*Functional organization of the dorsal spino-cerebellar tract in
the cat*
Acta Physiol Scand 38 1956 pp 53-75
- 12 O OSCARSSON
*Functional organization of the spino- and cuneocerebellar
tracts*
Physiol Rev 45 1956 pp 495-522
- 13 W T THACH
*Somatosensory receptive fields of single units in cat cerebellar
cortex*
J Neurophysiol 30 1967 pp 675-696
- 14 J S ALBUS
Control of a manipulator by a computer model of the cerebellum
PhD Thesis University of Maryland 1972