Biol. Cybern. 65, 267–279 (1991)

Biological Cybernetics

A neural network model of the cerebellar cortex performing dynamic associations

F. Chapeau-Blondeau[1] and G. Chauvet[1, 2]

[1] Institut de Biologie Théorique de l'Ouest Université d'Angers, 10 rue André Boquel, F-49100 Angers, France
[2] Department of Behavioral Neuroscience, University of Pittsburgh, Pittsburgh, PA 15260, USA

Received March 13, 1991/Accepted in revised form May 8, 1991

Abstract. The present paper proposes a model which applies formal neural network modeling techniques to construct a theoretical representation of the cerebellar cortex and its performances in motor control. A schema that makes explicit use of propagation delays of neural signals, is introduced to describe the ability to store temporal sequences of patterns in the Golgi-granule cell system. A perceptron association is then performed on these sequences of patterns by the Purkinje cell layer. The model conforms with important biological constraints, such as the known excitatory or inhibitory nature of the various synapses. Also, as suggested by experimental evidence, the synaptic plasticity underlying the learning ability of the model, is confined to the parallel fiber – Purkinje cell synapses, and takes place under the control of the climbing fibers. The result is a neural network model, constructed according to the anatomy of the cerebellar cortex, and capable of learning and retrieval of temporal sequences of patterns. It provides a framework to represent and interpret properties of learning and control of movements by the cerebellum, and to assess the capacity of formal neural network techniques for modeling of real neural systems.

1 Introduction

The cerebellar cortex is a biological system exhibiting both a structure and a general function which are relatively well known. From an anatomical standpoint, its structure shows great uniformity of constitution over its whole extent. From a physiological standpoint, it has been recognized that its general function is to play a role in motor control. This relative simplicity, both anatomically and physiologically, prompts us to look for a theoretical description of the neuronal processes that take place in the cerebellar cortex. The general problem raised by such a theoretical modeling of the cerebellar cortex is to determine an approach that allows us to capture the mechanisms of information processing which result in performances of learning, coordination, and/or control of movements.

Different theoretical models have been proposed for the cerebellar cortex in the past. Both Marr (1969) and Albus (1971) suggested a representation based on a specific assumption concerning the coding of information at the level of the mossy fibers and granule cells, and on an association mechanism for spatial patterns that is of a perceptron type. These models, however, do not address the temporal aspect of information processing which is necessary to account for properties of learning and coordination of movements.

Fujita (1982) introduced an adaptive filter model that relies mainly on control and optimization theories. This model exhibits temporal information processing abilities, but it fails to respect biological constraints. Moreover, it makes use of a learning mechanism based on a mean-square minimization whose biological plausibility is difficult to justify. Dunin-Barkowski and Larionova (1985) presented an application of concepts from information theory to evaluate the quantity of information that can be stored in the cerebellar cortex, depending on the properties and numbers of the different neurons. The work by Pellionisz and Llinas (1982), uses a new approach based on tensor calculus to describe neural processes. Although quite original and interesting, this work remains largely speculative.

In this paper, a new model is proposed, which shows temporal information processing abilities, and which seeks to closely respect both the known anatomy and physiology of the cerebellar cortex. Our approach utilizes and extends methods and techniques originally derived from the field of formal neural networks (Hopfield 1982; Rumelhart and McClelland 1986), and applies them to construct a theoretical model for a biological neural network. A schema is introduced, which makes explicit use of propagation delays of neural signals in a network, to represent a storage ability of temporal sequences of patterns in the Golgi-granule cell system. A perceptron association is then performed on these sequences of patterns by the Purkinje cell layer.

The model assigns a specific role to each of the different classes of neurons and fibers of the cerebellar cortex, and conforms with fundamental biological constraints, such as the known excitatory or inhibitory nature of the various synapses. As suggested by experimental evidence, and as also proposed by other investigators (Marr 1969; Albus 1971; Fujita 1982), the synaptic plasticity underlying the learning ability of our model, is confined exclusively to the parallel fiber – Purkinje cell synapses, and takes place under the control of the climbing fibers. The result is a neural network model, constructed according to the structure of the cerebellar cortex, and capable of learning and retrievel of temporal sequences of patterns. It provides a framework to represent and interpret performances of learning, coordination and control of movements of the cerebellum. It also gives an estimate of the capacity of formal neural network techniques for modeling of real neural systems.

2 Biological organization of the cerebellar cortex

Many details of the cerebellar anatomy and physiology still are largely unknown at the present time. The experimental information collected so far reveals an important anatomical regularity together with very complex functional properties. In this section, we briefly will review the main findings on the cerebellar cortex that are well accepted, and on which our model is based. The reference works are Eccles et al. (1967); Marr (1969); and Ito (1984).

The cerebellar cortex is a neural system (see Fig. 1) with two afferent pathways: the mossy fibers on the one hand, and the climbing fibers on the other hand. The efferent pathway is constituted exclusively by the Purkinje cell axons. The afferent pathways carry information from the cerebral cortex, or peripheral sensory organs; the Purkinje cells, in turn, project efferences into the motoneurons via cerebellar nuclei.

The incoming signals carried by the mossy fibers reach the glomeruli, which are complex synaptic structures distributing signals to multiple granule cell inputs. Granule cell axons form the parallel fibers, which can be viewed as a kind of data bus that conveys internal information throughout that cortex. Golgi cells are activated both by terminals of the parallel fibers, and by mossy fibers; they synapse back on granule cells and thus introduce a feedback loop. Purkinje cells receive inputs from parallel fibers, either through direct transmission, or through an indirect transmission via interneurons, namely the basket and the stellate cells (see Fig. 1). The Purkinje cells also receive inputs from the climbing fibers and those contacts appear to be one-to-one.

It sometimes is advantageous to consider the anatomical unit formed by the set of neurons and fibers which are directly or indirectly connected to a given Purkinje cell. Such a Purkinje unit in a human cerebellum contains approximately 80,000 granule cells. A given granule cell samples 3 to 5 mossy fibers, whereas

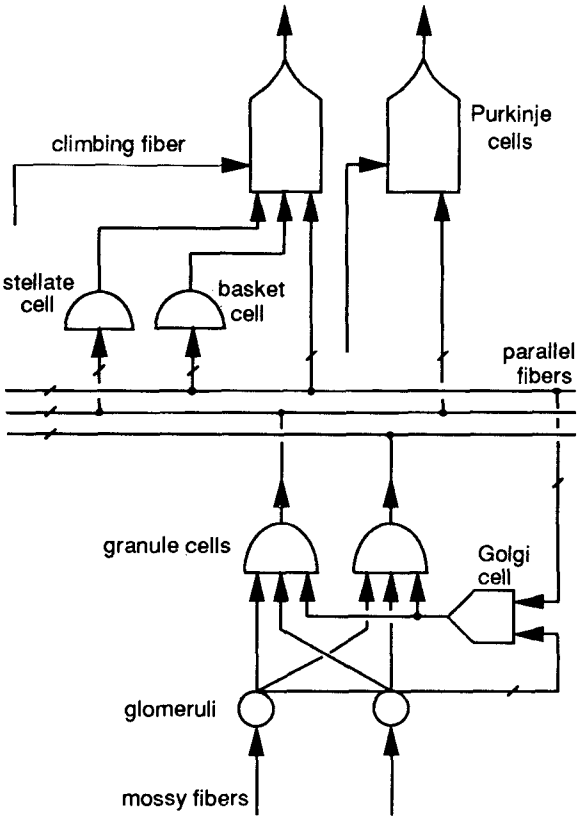

Fig. 1. Diagrammatic representation of the structure of the cerebellar cortex

a given mossy fiber makes contact with an average of 500 granule cells. There is one Golgi cell per 1 to 3 Purkinje units, and each Golgi cell synapses with approximately 5700 granule cells. A Purkinje cell receives inputs from 40 basket or stellate cells. It has been suggested (Chauvet 1986, 1988) that such a structural unit around a Purkinje cell may be considered also as a functional unit, in charge of the regulation of a given parameter in the coordination of movement. We also note the extreme divergence of the number of granule cells compared to the numbers of the other cells of the cerebellum. This unique feature of the granule cells needs to be accounted for in any model on cerebellar cortex.

It is generally assumed (see Eccles et al. 1967; Geiger 1977; Ito 1984) that in the neural network of the cerebellar cortex reside, at least for a part, the abilities of motor learning, motor coordination, and control of movement. However, little is known about the global information processing schemes, as well as the detailed mechanisms that underlie such performances.

3 A model for the Golgi-granule cell system

3.1 Structure and dynamics of the model

To represent the structure of the Golgi-granule cell system, we consider a neural network (see Fig. 2), consisting of a layer of a number N_{gr} of granule cells,

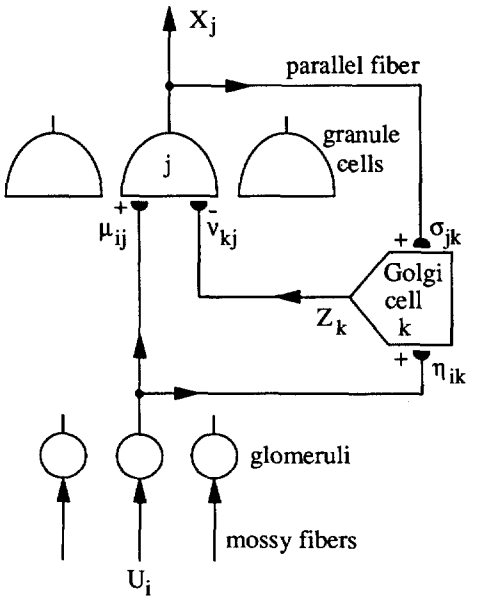

Fig. 2. Structure of the Golgi-granule cell system

referenced with index j ($j = 1$ to N_{gr}). This granule cell layer receives feedback from a set of a number N_{Go} of Golgi cells, referenced with index k ($k = 1$ to N_{Go}). Physiological data show that $N_{gr} \gg N_{Go}$. Signals from a set of N_m mossy fibers, referenced with index i ($i = 1$ to N_m), are distributed by the glomeruli to both granule and Golgi cell inputs. Granule cell outputs form the parallel fibers, which constitute inputs to the Golgi cells. Golgi cell outputs, in turn, are fed back on granule cell inputs.

The synaptic efficacies in this neural network are denoted as follows:
μ_{ij} for the synapse from mossy fiber i on granule cell j,
η_{ik} for the synapse from mossy fiber i on Golgi cell k,
σ_{jk} for the synapse from granule cell j on Golgi cell k,
v_{kj} for the synapse from Golgi cell k on granule cell j.
From physiological data (see Ito 1984) we know that the synapses represented by μ_{ij}, η_{ik} and σ_{jk} are excitatory, and thus will receive a positive sign in our model. On the other hand, the synapses v_{kj} are known to be inhibitory, and will be given a negative sign. After having presented the *structure*, we will now discuss the *dynamics* of the Golgi-granule cell system.

We note, at time t, $U_i(t)$ the signal conveyed by mossy fiber i, $X_j(t)$ the output signal of granule cell j and $Z_k(t)$ the output signal of Golgi cell k. To allow direct computer simulations, we chose a discrete-time representation for the evolution of the system, letting t take only integer values. The activities U_i, X_j and Z_k are quantized with values -1 or $+1$, and are interpreted as deviations from the mean activities of the neurons.

Let us assume that a Purkinje unit has a unique Golgi cell ($N_{Go} = k = 1$). This Golgi cell receives inputs from a large number of granule cells. These input signals have to travel along parallel fibers before they reach the Golgi cell; in this model, we shall explicitly take into account the traveling times of these signals along the parallel fibers.

For that purpose, we assume that the set of granule cells connected to the Golgi cell can be divided into a number N_{clas} of classes, which each contains a number N_{gpc} of granule cells. Thus, we assume that $N_{gr} = N_{clas} \cdot N_{gpc}$. These classes of granule cells are given a rank, ranging from rank 1 to rank N_{clas}. The rank of a given class expresses how distant this class is in spatial relation to the Golgi cells. So, if we consider the signals $X_j(t)$, present at time t on the granule cell outputs, we suppose that only the output signal or those granule cells belonging to class 1 can reach the Golgi cell at the same time t. The output signal from the granule cells of class 2 will reach the Golgi cell at time $t+1$, and so on, up to the output signal from the granule cells of class N_{clas}, which will reach the Golgi cell at time $t + (N_{clas} - 1)$.

For the calculation of the soma potential of the Golgi cell, we also have to include contribution of the input signals U_i coming from the mossy fibers. For the operation described in this model, we consider that these signals U_i are constant in time. We thus define the soma potential $V_1^{Go}(t)$ of the Golgi cell corresponding to $k = 1$, at time t, as:

$$V_1^{Go}(t) = \frac{1}{N_{gr}} \sum_{c=0}^{N_{clas}-1} \sum_{j=cN_{gpc}+1}^{(c+1)N_{gpc}} \sigma_{j1} X_j(t-c) + \frac{1}{N_m} \sum_{i=1}^{N_m} \eta_{i1} U_i . \quad (1)$$

We then express the output $Z_1(t+1)$ of the Golgi cell 1 at time $t+1$ as a non-linear function of its soma potential:

$$Z_1(t+1) = \mathrm{sgn}(V_1^{Go}(t)) . \quad (2)$$

Now, to express the soma potentials of the granule cells, we also take into account the propagation delays of the signals coming from the Golgi cell. In the same way, the output signal of the Golgi cell at time t, will reach a given granule cell at a later time, depending on the class to which this granule cell belongs. After also including the mossy fiber inputs, we express the soma potential $V_j^{gr}(t)$ of the granule cell j, at time t, as:

$$V_j^{gr}(t) = v_{1j} Z_1(t-c) + \frac{1}{N_m} \sum_{i=1}^{N_m} \mu_{ij} U_i,$$

$$\text{when} \quad j \in [cN_{gpc}+1, (c+1)N_{gpc}] ; \quad (3)$$

$$\text{for } c = 0 \quad \text{to} \quad N_{clas} - 1.$$

And the output $X_j(t+1)$ of granule cell j, at time $t+1$, is:

$$X_j(t+1) = \mathrm{sgn}(V_j^{gr}(t)) . \quad (4)$$

In the definitions (1) and (3) of the soma potentials, we have divided each partial sum that corresponds to the contribution of each category of cells, by the number of cells in this category. This operation provides all categories of cells with a comparable relative importance in the dynamics of the network. Failing to do

such a normalization process of the activities of the different categories of neurons would imply that only the most numerous category of cells plays a significant role in the dynamics of the system.

To specify the values of the synaptic efficacies of this neural network, we assumed, as it has been done by other investigators (Marr 1969; Albus 1971; Fujita 1982), that no synaptic plasticity occurs in the Golgi-granule cell system. Therefore, each value of the synaptic efficacies of this system is fixed and assigned a sign according to the known exitatory or inhibitory nature of the synapse.

Once the values for the synaptic efficacies have been specified, (1), (2), (3) and (4) determine the dynamics of the Golgi-granule cell system. For operation of the network, a constant input pattern **U** is presented to the mossy fibers. As a result, the output pattern $\mathbf{X}(t)$ formed by the granule cell outputs at time t, shows a complex evolution as time proceeds, due to the feedback lop at the Golgi-granule cell level. Thus, for a given constant input pattern **U**, the iteration of the network dynamics generates a temporal sequence of output patterns $\mathbf{X}(t)$, which is labelled by the input pattern **U**. The presentation of a different input pattern **U** will generate a different sequence of output patterns $\mathbf{X}(t)$.

This important property of sequence generation of the neural network is a consequence of both the presence of a feedback loop and the intervention of propagation delays in the system. The feedback loop is necessary to endow the network with an intrinsic dynamic behavior, instead of a static one-to-one input–output association behavior, as has been described in earlier models (Marr 1969; Albus 1971). The propagation delays are useful to provide sufficiently rich dynamic evolutions, in the presence of a very absorbant feedback loop, as revealed by anatomical data (a large number of granule cells fed back through a single Golgi cell). With such an absorbing feedback, simulations have shown that no interesting evolutions of the system can be obtained if propagation delays between the incoming signals on a Golgi cell are not taken into account, because in such a case the dynamic trajectories would be reduced to fixed points or two-state trajectories most of the time (Chapeau-Blondeau et al. 1989).

3.2 Simulation results

Equations (1), (2), (3) and (4) have been used for computer simulations of the time evolution of the Golgi-granule cell system. These simulations allowed us to visualize the temporal sequences generated by the network in various situations, and to study the influence on these sequences of the four principal parameters on which they depend:

i) the initial values assigned to the patterns **X** and **Z**;
ii) the numbers of the different cells in the network, especially N_{gpc} and N_{clas};
iii) the particular set of values chosen for the synaptic efficacies; and
iv) the constant input pattern **U** that is presented to the network.

For the simulations, the values of the synaptic efficacies were chosen randomly, with a uniform probability, between 0 and 1 for the excitatory synapses μ_{ij}, η_{ik} and σ_{jk}, and between -1 and 0 for the inhibitory synapses ν_{kj}. The influence of the procedure for the choice of synaptic values will be discussed later.

Within the framework of this discrete representation, only a finite number of output states **X** are accessible to the network. Moreover, because the dynamics of the network is purely deterministic, the time trajectory of $\mathbf{X}(t)$ always reaches, after a finite transient path, a periodic attractor or cycle. This limit cycle may reduce, in some cases, to a single state or fixed point, but this outcome becomes very unlikely as the number of cells in the system becomes large.

Stable rest state

A case in which the dynamics always was found to converge onto a fixed point, was the situation in which is presented to the network an input pattern **U** that contains only -1's. In such case, the network rapidly reaches a fixed output state $\mathbf{X}_s$ which depends on the set of values of the synaptic efficacies, but which is found independent of the initial values of the patterns **X** and **Z**. Such a fixed point $\mathbf{X}_s$ represents the stable state at which the network arrives in response to a complete deficit of input activities, represented by a pattern **U** containing only -1's. This stable state can be viewed as the rest state of the network in the absence of activity on the mossy fibers. In general, this stable state $\mathbf{X}_s$ is a N_{gr}-component vector that contains both -1's and $+1$'s. It is this vector $\mathbf{X}_s$, and the associated vector $\mathbf{Z}_s$, present in the stable state, that we took as initial quantities for the iteration of the dynamics of the network in response to an arbitrary input pattern **U**.

Influence of the numbers of cells

When iterating the dynamics, starting from the stable state of the network, it was observed that both the transient path and the limit cycle of the output trajectory increase in length, as the numbers of the different cells increase.

Let us consider, for example, a small network with one Golgi cell associated to $N_m = 4$ mossy fibers, and $N_{gr} = 20$ granule cells (split into $N_{clas} = 10$ and $N_{gpc} = 2$). In such a network, $2^{N_m} = 16$ different input patterns **U** can be presented, and a very large number $2^{N_{gr}} \approx 10^6$ of different output patterns **X** can result. For a given set of synaptic efficacies, the numerical simulations of the dynamics showed the various temporal sequences that can be generated by the network in response to different input patterns **U**. Depending on **U**, the output sequences observed were formed typically of a transient path, of variable length, containing 10 to 100 states, followed by a limit cycle that also was variable in length, and contained a few tens of states. These "typical" properties were maintained upon changing the values for the synaptic efficacies; however, the particular sequence generated in response to a given **U** generally differed.

As the numbers N_{gr} and N_m increase, the trajectories of $\mathbf{X}(t)$ become richer and richer: their variety, transient part, and limit cycle all were found to increase. For a sufficiently large number of granule cells ($N_{gr} = 200$, $N_{gpc} = 2$ for example), both the transient path and limit cycle are too long to be identified separately, and the trajectory appears as a non-periodic sequence of patterns. In biological neural networks, this type of situation can be expected to apply, due to the large number of cells (especially of granule cells).

In the model, the exact decomposition of the product $N_{clas} \cdot N_{gpc}$ in which N_{gr} is partitioned, does not play a critical role, and the characteristic performances of the network are maintained provided N_{clas} is sufficiently large (to preserve notable contribution by the propagation delays). When the number N_{clas} was reduced to 1, the output trajectories shrunk to fixed points or two-state trajectories.

Influence of the synaptic efficacies

Simulations also revealed that the procedure for the choice of the synaptic efficacies does not alter the main properties of the model, provided that procedure preserves the prescribed synaptic signs. Specifically, we also tested the procedure for assigning values to the synaptic efficacies, which consists in forcing the excitatory synapses with value $+1$, and the inhibitory synapses with value -1. This choice did not alter the ability of the network to generate temporal sequences labelled by an input pattern **U**. However, in such extreme case, part of the discrimination power of the system was lost, because all input patterns **U** that differed only by circular permutations of their components, were associated with the same sequence of output patterns. A random distribution of synaptic efficacies around a center value seems a plausible feature from a biological standpoint, and yields greater discrimination power of the system. Specifically, it seems to enable it, in some sense, to distinguish between individual mossy fibers.

We also investigated the effect of dilution among the synapses of our network model. Starting with a set of random synapses in the network, we then forced a fraction of them to zero, by random choice out of the four synaptic classes.

For small degrees of dilution (1% of the synapses or less) no significant alterations were induced in generated output sequences $\mathbf{X}(t)$ of reference length 100 time steps. In particular, the limit cycles were not altered in general. From a biological standpoint, such small degrees of dilution of the synapses would result from natural synapse breakdowns in the system. If dilution adds up with time to become more important (superior to a few percents), this process can be expected to be slow compared to the sequence generation time scale. Such a quasi static drift of the synapses can be compensated for by permanent learning occuring over long time scales, and does not interfere significantly with the short-time scale operation of sequence generation.

From the anatomy of the cerebellar cortex, we know that, strictly, there does not exist a synapse connecting any possible mossy fiber-granule cell pair. To represent and analyze this absence of full connectivity in the structure of the system, we diluted important fractions of the synapses (10% or more) and examined how the performance of the network model was affected. It appeared that a large degree of dilution can be tolerated without alteration of the global sequence generation ability of the network. The main effect of large synaptic dilution was to reduce the lengths of both the transient path and the limit cycle in the output trajectory of the network. For instance, for the particular network evoked above with $N_{gr} = 200$, with a fraction of 90% of the synapses set to zero, sequence generation was still observed, with trajectories showing transient paths of few tens of states and limit cycles reduced to fixed points. This suggests that dilution of synapses can provide a means for the network to control the length of the generated sequences.

Perturbation resistance

The network, when performing an association between one input pattern **U** and a sequence of output patterns **X**, can tolerate a certain degree of distorsion of the input pattern, and still reproduce reliably the output sequence. Two types of distorsion of the inputs were envisaged.

One type was to perturbe each signal U_i by addition of a random noise of small magnitude, and then examine how the output sequence was affected. For an output sequence of a reference length of 100 time steps, we have computed a mean Hamming distance between the patterns **X** that correspond to one another in the output sequences generated in the absence and in the presence of noise (the Hamming distances were normalized to 1 by a division by N_{gr}). This mean distance was then averaged over 1000 sequences of length 100 randomly picked, to yield an average separation corresponding to a given noise level. This average separation is plotted in Fig. 3 as a function of the magnitude of a uniform noise applied to the inputs U_i. These results were obtained for a network with $N_m = 100$ mossy fibers and $N_{gr} = 100$ granule cells.

For instance, for a noise level of 0.1 in Fig. 3, an average separation of 0.014 means that in a sequence of 100 time steps, the output of a given granule cell was in the wrong state of activity in 1.4 steps on average. For a noise level of the order of 0.05 or below, there is, on average, less than one error in the output of a granule cell when generating a sequence of length 100, which gives good probability to reliably reproduce the sequence. With this type of perturbation, the noise resistance was found to increase with the number of granule cells and mossy fibers in the system. This can be explained by the fact that noise tends to average out when weighted sums of activities are performed by each cell evaluating its soma potential through (1) or (3).

Another type of perturbation of the inputs was envisaged, in which, for a small fraction of the signals U_i of an input pattern **U**, the state of activity was simply reversed. The "perturbed" output sequence gen-

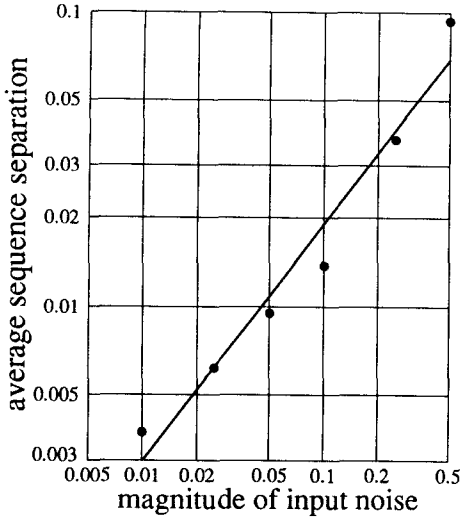

Fig. 3. Average separation of a perturbed and an unperturbed output sequence, as a function of the magnitude of a continuous noise applied to input signals U_i, and obtained for a network with $N_m = 100$ mossy fibers and $N_{gr} = 100$ granule cells

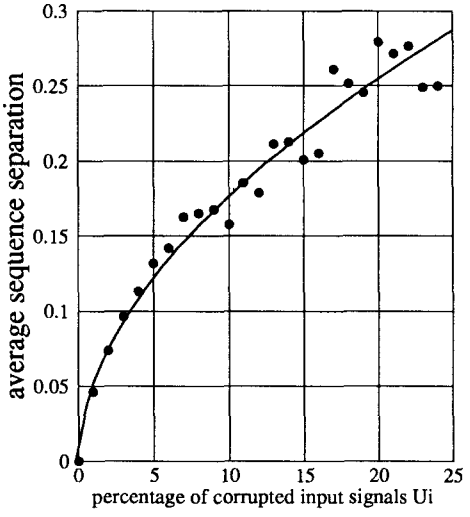

Fig. 4. Average separation of a perturbed and an unperturbed output sequence, as a function of the percentage of corrupted (reversed) input signals U_i, and obtained for a network with $N_m = 100$ mossy fibers and $N_{gr} = 100$ granule cells

erated under these conditions was then compared to the unperturbed sequence with no reversion of the inputs. As before, we computed the mean normalized Hamming distance between the corresponding patterns in a perturbed and unperturbed output sequences of reference length 100. This mean distance was then averaged over 1000 different output sequences, to yield the average separation occuring in the presence of a given fraction of reversed inputs. Figure 4 shows such an average separation as a function of the percentage of corrupted (reversed) input signals U_i, obtained for a network with $N_m = 100$ mossy fibers and $N_{gr} = 100$ granule cells.

Here, for a fraction of 5% of corrupted U_i in Fig. 4, an average separation of 0.13 means that in a sequence of 100 time steps, the output of a given granule cell was in the wrong state of activity in 13 steps on average. With this type of perturbation, the noise resistance associated to a given fraction of corrupted inputs, was not improved significantly by increasing the number of mossy fibers or granule cells in the system. The comparison of Fig. 4 with Fig. 3 shows that the perturbation resistance of the system in presence of input reversion is weaker than in presence of continuous noise. Only for less than one percent of reversed inputs, may we expect to have, on average, less than one error in the output of a granule cell when generating a sequence of length 100. Such a behavior in response to reversions in the input pattern implies, in counterpart, that the system is capable of learning distinctively a larger number of output sequences. We may not a priori expect, from a neural network memorizing dynamical sequences, the same degree of generalization capability as exhibited by attractor neural networks memorizing static patterns. The main task of the cerebellar cortex in the framework of the present model is to bring dynamical treatment abilities. Error resistance may be provided by other neural circuits associated with the cerebellar cortex, as those located in the cerebellar nuclei, and behaving as attractor neural networks, to ensure for instance a correct input pattern **U**. Then, given this **U**, the cerebellar cortex system can perform reliable sequence generation, even in the presence of continuous noise superimposed to the inputs.

Decorrelation of the patterns

Another property of the neural network is the high level of decorrelation it provides between the output patterns of a sequence. We define the correlation of two patterns as their normalized dot product $\langle \mathbf{X}^p \cdot \mathbf{X}^q \rangle = (1/N_{gr}) \Sigma_j X_j^p X_j^q$. Then, any two patterns of a sequence generally show a low correlation (a good decorrelation), except of course for two occurences of the same patterns in a periodic cycle (see Sect. 5). This decorrelation property makes the output patterns well suited for efficient use in a perceptron association scheme (see below). The correlation decreases when the numbers of cells in the network increase. From numerical simulations it comes out that, for a given network, both the length of the sequences generated before cycling, and the correlations among the patterns in these sequences, depend on the values of the synaptic efficacies. With the network dynamics chosen here, there certainly exist optimal sets of synaptic efficacies that lead to sequences of maximal length or minimal correlation. However, we did not find any obvious means to theoretically derive those optimal sets. Instead of solving an optimization problem, another way to attain to long sequences of decorrelated output patterns is to increase the number of cells in the network. Considering the very large number of granule cells existing in the cerebellar cortex, this could be the solution implemented in biological systems.

4 The Purkinje cell layer

4.1 A perceptron model

The model proposed here for the Golgi-granule cell system exhibits properties of sequence generation. A temporal sequence of output patterns is produced as a response to a constant input pattern **U**. The output sequence, labelled by input pattern **U**, is formed with output patterns that cannot be specified, but which are imposed by the network characteristics. The capacity of granting significance to these sequences of nonspecific patterns, is a role which can be assigned to the Purkinje cell layer. For that purpose we assume in our model, that the Purkinje cell layer acts as a multiple perceptron (Minsky and Papert 1969), which transforms sequences of non-specific patterns generated by the Golgi-granule cell system, into sequences of patterns that encode actual commands for the motor system.

Such perceptron behavior by the cerebellar cortex has been proposed by Albus (1971). However, in previous models, the perceptron was assumed to perform an association of a static type, between one mossy fiber input pattern and one output pattern on the Purkinje cell axons. In contrast, our model results in a dymanic association between one mossy fiber input pattern and a sequence of output patterns on the Purkinje cell axons, due to the sequence generation property of the Golgi-granule cell system.

A given Purkinje cell, referenced with index m, can be considered as the output unit of a simple perceptron. Input signals for this perceptron are constituted by the output activities **X** of the Golgi-granule cell system. Input signal X_j is applied to Purkinje cell m through a synaptic efficacy w_{jm}. The output signal of Purkinje cell m will be noted Y_m.

Let us denote by $\mathbf{X}^p$ (where p is a variable index) the different patterns appearing in an input sequence. The role of the Purkinje cell layer is to map this sequence of input patterns $\mathbf{X}^p$, into a sequence of corresponding output patterns $\mathbf{Y}^p$. The high level of decorrelation that can be obtained by the Golgi-granule cell system between the different patterns $\mathbf{X}^p$ of a sequence, is a favorable feature which yields a high discrimination power in a perceptron association mechanism.

To achieve the desired mapping, the presentation of a given pattern $\mathbf{X}^p$ during a learning phase, produces a change $\Delta^p w_{jm}$ in the synpatic efficacy w_{jm}, which can be expressed, according to classical perceptron learning rule, by:

$$\Delta^p w_{jm} = X_j^p Y_m^p . \tag{5}$$

To make this learning scheme realizable in biological systems, there is a need for a mechanism that informs the system of the desired output Y_m^p which has to be associated with each input X_j^p. Here, we resort to an assumption, that has been introduced in modeling previously (Marr 1969; Albus 1971; Fujita 1982), and has been demonstrated experimentally (see Ito 1984): learning at the Purkinje cell level is performed under the control of the climbing fiber. We denote by C_m the signal carried by the climbing fiber of Purkinje cell m, and we suppose that when learning takes place, the specific output signal Y_m^p, which has to be associated with a given input signal X_j^p, is presented to the Purkinje cell under the form of a specific climbing fiber signal C_m^p. The learning rule (5) can thus be rewritten as:

$$\Delta^p w_{jm} = X_j^p C_m^p . \tag{6}$$

Assuming learning starts from a tabula rasa (i.e., the synapses all are zero before they learn), after presentation of a sequence to be learned, the synaptic efficacy w_{jm} has taken the value:

$$w_{jm} = \sum_p \Delta^p w_{jm} = \sum_p X_j^p C_m^p . \tag{7}$$

It would be desirable that the system be able to distinguish by itself between *learning phases*, which involve the learning of novel movements by modification of synaptic efficacies, and *retrieval phases*, which involve the generation of movements without modification of the synapses. This performance can be achieved by allowing a climbing fiber to be in an active state encoded by the value $C_m^p = +1$, or in an inactive state, encoded by the value $C_m^p = 0$. By doing this, the climbing fiber can be assigned the double role of a teacher signal, and of switching between learning phases (when $C_m^p = +1$) and retrievel phases (when $C_m^p = 0$). Before demonstrating this feature of the system, we point out that the role of the climbing fiber, as described here, also can be interpreted as that of an error signal. Such role has been assigned to the climbing fiber by other investigators (Fujita 1982).

With our representation of climbing fiber activity, it is clear from (6) that the variation of the synaptic efficacy only takes place when $C_m^p = +1$ and vanishes when $C_m^p = 0$, thus implementing the switching mechanism between learning phases and retrieval phases of the system.

Now for the retrieval operation, we define the soma potential V_m^{Pu} of Purkinje cell m when an arbitrary input pattern **X** is presented, as:

$$V_m^{Pu} = \frac{1}{N_{gr}} \sum_{j=1}^{N_{gr}} w_{jm} X_j . \tag{8}$$

The resulting value for the output Y_m of Purkinje cell m will be defined as:

$$Y_m = \mathrm{sgn}(V_m^{Pu} - \theta) , \tag{9}$$

where θ is a threshold we suppose to exist for the activation of Purkinje cells.

If we substitute (7) in (8), and make the suitable arrangements, we can express the soma potential as:

$$V_m^{Pu} = \sum_p C_m^p < \mathbf{X} \cdot \mathbf{X}^p > , \tag{10}$$

where $<\mathbf{X} \cdot \mathbf{X}^p>$ stands for the normalized dot product of presented pattern **X** and learned pattern $\mathbf{X}^p$.

Now let us assume that the presented pattern **X** is equal or close to a particular learned pattern, say $\mathbf{X}^{p_0}$. Then we shall have:

$$<\mathbf{X} \cdot \mathbf{X}^p> \approx 1 \qquad \text{for} \qquad p = p_0 ,$$

and, for a sufficient level of decorrelation between learned patterns:

$$\langle \mathbf{X} \cdot \mathbf{X}^p \rangle \approx 0 \qquad \text{for} \qquad p \neq p_0 .$$

The soma potential thus reduces to:

$$V_m^{Pu} = C_m^{P_0} .$$

Therefore, it is the value of the teacher signal $C_m^{P_0}$ presented during learning, which will determine, through (9), the output value Y_m of Purkinje cell m during the retrieval operation. The generation of an active output $Y_m = +1$ when the teacher signal is $+1$, and an inactive output $Y_m = -1$ when the teacher signal is 0, will be most efficiently achieved when the Purkinje cell threshold is set to the mean value of climbing fiber activity, i.e., when $\theta = 0.5$.

As stated above, the perceptron learning described here, can also be interpreted as governed by an error signal rather than by a teacher signal. For a Purkinje cell m, no learning takes place if the climbing fiber is not activated. And if no learning has occurred, the output of Purkinje cell m will remain in a low-activity state coded by $Y_m = -1$. If a situation requires the output Y_m to be in a state of high activity coded by $+1$, then only in this case the climbing fiber is activated, leading the Purkinje cell to learn to generate a high-activity output as a response to its current input. Thus, only conditions where a high-activity output $Y_m = +1$ is required have to be explicitly learned, and are effectively learned in a mechanism triggered by the activation of the climbing fiber error signal. Conditions where a low-activity output $Y_m = -1$ is required correspond to a passive state of the cell, and do not have to be explicitly learned. From a biological standpoint, this learning scheme assigns to the climbing fiber the role of an error signal, rather than that of a teacher signal. The realizability of this scheme does not require the pre-existence in the system of the activity patterns to be learned, but only the ability to learn to generate high-activity outputs in response to the activation of the climbing fibers.

4.2 Sign constraint for Purkinje cell synapses

Experimental data show that the synapses from parallel fibers on Purkinje cells are excitatory. With learning rule (6), the synaptic efficacies w_{jm} can be either positive or negative, because the X_j^p's can be $+1$ or -1. Assignment of physiologically correct signs to synaptic efficacies can be achieved by consideration of the two possible pathways from parallel fibers on Purkinje cells: a direct pathway that involves excitatory synapses, and an indirect one that involves inhibitory interneurons, the basket and stellate cells. We shall simplify the complex connectivity by assuming that the connection of an interneuron is as depicted on Fig. 5. Next, as done by Fujita (1982), we split the synaptic efficacy w_{jm} into two parts, $w_{jm} = w_{jm}^- + w_{jm}^+$, where w_{jm}^- is the inhibitory synapse of the indirect pathway, and w_{jm}^+ is the excitatory synapse of the direct pathway. We assume

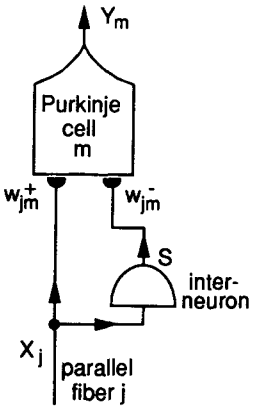

Fig. 5. Double-pathway connection from parallel fiber j to Purkinje cell m

that the interneuron of Fig. 5 performs an input–output transformation which is the identity transformation expressable as:

$$S = \mathrm{sgn}(X_j) = X_j .$$

Such a mechanism allows us to preserve the correct signs for the different synapses, with the learning scheme of (6). When the learning process (6) builds up a positive value w_{jm}, its value is assigned to w_{jm}^+ whilst w_{jm}^- stays at a zero value; the opposite takes place when learning yields a negative w_{jm}.

We note, however, that such a double-pathway connection is unlikely to exist for all possible parallel fiber-Purkinje cell pairs. This would require a number of interneurons as large as the number of granule cells times the number of Purkinje cells. It is known that the number of interneurons in the cerebellum is only intermediate between that of the Purkinje cells and that of the granule cells.

In any case, a negative-to-positive range for the value of the synaptic contacts for each possible parallel fiber-Pukinje cell pair is not necessary to preserve correct perceptron behavior of the system. *(i)* It is well known (see for instance Shinomoto 1987) that a certain amount of dilution in the synapses of a neural network does not suppress its global pattern storage ability. Thus, the fraction of parallel fiber-Purkinje cell synapses that do not possess the double-pathway connection, could be viewed as included among the diluted synapses with no dominant role in the system. *(ii)* It recently has been shown (Amit et al. 1989) that the perceptron behavior can subsist with any arbitrary set of signs imposed on the synapses, even with purely excitatory synapses. As in the case of dilution, such constraints lead to a reduction of the storage capacity of the network. Storage capacity can be restored, however, by increasing the number of input units (i.e., the granule cells) of the perceptron. This could provide an explanation for the very large number of granule cells in the cerebellar cortex.

4.3 Numerical simulations

The perceptron operation has been tested numerically with learning rule (6). Three different possibilities were examined to ensure physiologically correct signs to the synaptic efficacies. First, we allowed both signs to be available for all synapses of the perceptron, assuming a double-pathway connection for all possible synaptic contacts between a parallel fiber and a Purkinje cell. Second, we mixed synapses with two possible signs with positive only synapses. Third, we assumed only positive synapses. Learning for a positive synapse is still performed according to rule (6), except when the modification prescribed by this rule would make the synapse negative, in which case such modification would be not performed.

With all three possibilities, a correct associative perceptron behaviour was observed.

An important aspect to verify was the high probability for patterns to be correctly associated by the perceptron. It is known that correct classification of a set of patterns by a perceptron requires fulfillment of the condition of linear separability of the patterns (Minsky and Papert 1969). It also is known that correct classification of a given set of patterns can always be attained by addition of further additional units (generally, hidden units) to the perceptron. In the model studied here, the role assigned to the perceptron is that of learning and retrieval of a set of output patterns specified by the climbing fibers. This is performed by associating these sequences of specific output patterns to sequences of non-specific input patterns. As these input patterns (generated by the Golgi-granule cell system) cannot be specified, they can be considered as randomly appearing. We thus have studied the probability for a given set of input patterns, picked at random, to be correctly mapped onto a given set of output patterns by the perceptron. Typical results on such operation are shown in Fig. 6, which represents the probability of correct association as a function of the number of granule cells (input units of the simple perceptron), in different conditions.

Figure 6a represents the probability of correct embedding of a given set of 5 input–output pairs of patterns. This probability was evaluated as the frequency of correct embedding calculated by numerical simulations for a given number of granule cells, with 5000 sets of five input–output pairs randomly selected. In this case, the synaptic efficacies of the simple perceptron performing the association were not sign-constrained (i.e., they could be positive or negative). Figure 6b represents the probability of correct embedding of a given set of 10 input-output pairs, using the same method. Figure 6c shows the same probability as in Fig. 6a, except that this time the synapses of the perceptron were constrained to be positive.

The results of Fig. 6 show that for a given value of the number of granule cells N_{gr}, the probability of correct embedding of a given set of patterns varies with the number of patterns to be stored and with the constraints imposed onto the network. Nevertheless,

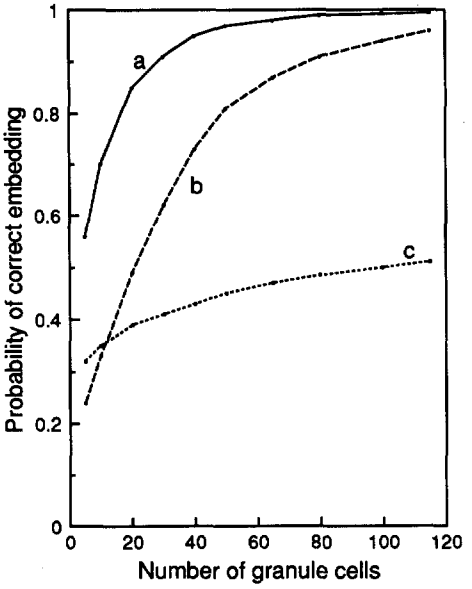

Fig. 6. Probability of correct embedding of N_p input patterns randomly selected, as a function of the number of granule cells of the network. *a)* for $N_p = 5$ and *b)* for $N_p = 10$, with no sign constraints on the synapses. *c)* for $N_p = 5$ with sign-constrained synapses

under all conditions, this probability is an increasing function of the number of granule cells N_{gr}. The simulations suggest that this probability can be made very close to unity by increasing the value of N_{gr}. (For the constrained case depicted in Fig. 6c, the largest value $N_{gr} = 2500$ tested gave a probability of 0.73). We can thus deduce that any arbitrary set of patterns can be stored in the perceptron, even with sign-constrained synapses (see Amit et al. 1989, for exact proof), provided the number of granule cells is sufficiently large. If we remember that, in the cerebellar cortex, an average number of 80,000 granule cells are available for one Purkinje cell, we can infer that with such a mechanism, the system will be capable of storing a very large number of sequences of patterns, with a good probability of correctness.

5 Simulation of the complete system

When the two levels of the model are joined together, the result is a neural network that can learn and retrieve temporal sequences of specified output patterns. The control of the length of the output sequences on the Purkinje cell axons can be carried out by the system in different ways. A simple way is to present on the mossy fiber inputs a complete deficit of activity, coded by a pattern $\mathbf{U}$ containing only -1's. After a short transient, the system will then respond by converging to a stable point that corresponds to a fixed intermediate pattern $\mathbf{X}_s$ mapped on to a fixed output pattern $\mathbf{Y}_s$. If one wants to suppress on the output trajectory $\mathbf{Y}(t)$ the transient phase that precedes stabilization, one can simply map

Table 1. Time evolution of the patterns in a retrieval phase of a neural network with 20 granule cells. *First column*: applied input patterns **U**, *second column*: intermediate patterns **X**, *third column*: learned output patterns **Y**, *fourth column*: normalized dot product of each pattern **X** with the fourth pattern $\mathbf{X}_4$. (In **U** and **Y** a "0" means a -1 activity state)

U	X	Y	$<\mathbf{X}\cdot\mathbf{X}_4>$
00000000	653193	00	−0.2
10101010	436613	01	0.1
10101010	874603	10	−0.1
10101010	035647	11	1
00000000	653189	00	−0.2
00000000	653193	00	−0.2

the few intermediate patterns **X** of the transient on to the desired stable output $\mathbf{X}_s$.

The following examples we present are illustrations of the abilities of the complete system to deal with temporal sequences of patterns.

The first example uses a network with 8 mossy fibers, 20 granule cells (split into $N_{clas}=5$ and $N_{gpc}=4$), 1 Golgi cell, and 2 Purkinje cells. The network is trained to be in a stable output state coded as $\mathbf{Y}=--$, when input pattern $\mathbf{U}=--------$ is presented; to generate the output sequence formed with successive patterns $\mathbf{Y}=-+$, $\mathbf{Y}=+-$, $\mathbf{Y}=++$, when the input pattern $\mathbf{U}=+-+-+-+-$ is presented; and then to stop and return to stable state $\mathbf{Y}=--$. Table 1 gives, at each iteration of a retrieval phase, the time evolution of the output pattern **Y** in response to the input pattern **U** which is applied. Also given are the nonspecific patterns **X** which serve for an intermediate coding of the sequences of patterns. Values of the normalized dot product of the different patterns **X** with a given one among them, are also presented as an indicative estimation of their decorrelation. In Table 1, we used a binary representation of patterns **U** and **Y** (with "0" for a -1 activity state, and "1" for a $+1$ activity state), and a decimal representation for patterns **X**. Figure 7 shows the time signal which can be obtained by plotting the decimal value of output pattern **Y**, when different successive values of input pattern **U** are applied.

The second example uses a network with 8 mossy fibers, 80 granule cells (split into $N_{clas}=20$ and $N_{gpc}=4$), 1 Golgi cell, and 3 Purkinje cells. The network is trained to memorize two different output sequences this time, instead of one. It is in a stable output state that we now choose to code as $\mathbf{Y}=+--$, when input pattern $\mathbf{U}=---------$ is applied; it retrieves the first output sequence if input pattern $\mathbf{U}=+-+-+-+-$ is presented, and the second if $\mathbf{U}=-+-+-+-+$ is presented. Table 2 shows the corresponding data. Figure 8 presents the time signal obtained in the same way as for Fig. 7. This type of signals can be viewed as command signals for the temporal control of a given parameter in a movement, such as the angle of a joint or the tension of a muscle.

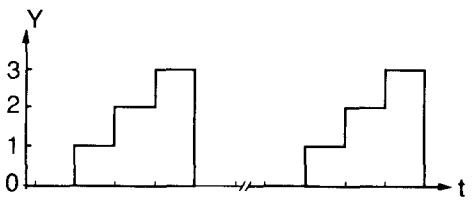

Fig. 7. A temporal representation of a sequence memorized by a network with 20 granule cells

Table 2. Time evolution of the patterns in a retrieval phase of a neural network with 80 granule cells. *First column*: applied input patterns **U**, *second column*: learned output patterns **Y**, *third column*: normalized dot product of each intermediate pattern **X** with the second pattern $\mathbf{X}_2$. (In **U** and **Y** a "0" means a -1 activity state)

U	Y	$<\mathbf{X}\cdot\mathbf{X}_2>$
00000000	100	−0.075
10101010	101	1
10101010	111	−0.05
10101010	101	−0.075
00000000	100	0.125
00000000	100	−0.075
00000000	100	−0.075
00000000	100	−0.75
00000000	100	−0.075
01010101	001	0.025
01010101	010	0.0
01010101	001	0.025
00000000	100	−0.025
00000000	100	−0.075

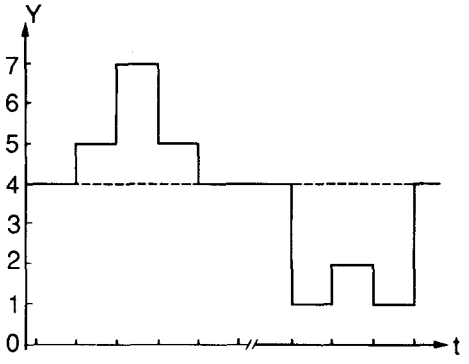

Fig. 8. A temporal representation of two different sequences memorized by a network with 80 granule cells

These examples illustrate how it is possible to handle sequences of patterns of various forms, even with small networks. It is clear that with networks of the size of biological networks, which contain a very large number of neurons, important storage capacity (in terms of number and length of the sequences) can be accomplished. We did not investigate the optimal storage capacity as a function of the number of cells. The important feature we wanted to exemplify is the ability of a neural network, constructed according to the

anatomy of the cerebellar cortex, to deal with temporal sequences of patterns.

6 Discussion of the model

6.1 Neural network theory standpoint

We have introduced in this paper a biologically constrained neural network model, to represent the neuronal system of the cerebellar cortex. This model is able to learn and retrieve temporal sequences of patterns.

The architecture of this network model closely adheres to the anatomy of the cerebellar cortex. Its structure is arranged in three layers (glomeruli, granule cells, Purkinje cells) which include a feedback loop from the granule cell layer through Golgi cells, back to the granule cells. It is to note however that we are not dealing here with a classical layered network, because of the presence of the feedback. The structure neither represents a fully recurrent network, because of the strictly limited scope of the feedback. We believe that such intermediate network architectures, based on biological grounds, can provide fruitful frameworks to generate new capabilities for neural network models.

In the network model considered here, we explicitly took into account the propagation delays of a certain class of neural signals in the net. Such assignment of a role to propagation delays is a feature seldom considered in modeling with formal neural networks; yet, it is likely to be present in biological neural networks. Incorporating such delays into the dynamics also gives rise to new and interesting network behaviors. In the present model, propagation delays intervening in a feedback loop of a very absorbing type prevent the dynamics of the net from being blocked too easily in fixed points.

It is also interesting to emphasize that the neural network model described here possesses different classes of synapses, most of which are not modifiable. This demonstrates that neural networks with non-modifiable synapses are not necessarily confined to trivial behaviors, and may exhibit fundamental properties, such as the sequence generation ability of the Golgi-granule cell system.

The present neural network model, offers a scheme to learn temporal sequences of patterns, and to retrieve them upon presentation of the input label to which they were associated. This constitutes a proposal, among others recently introduced (Sompolinsky and Kanter 1986; Dehaene et al. 1987; Kuhn et al. 1989; Massone and Bizzi 1989; Pearlmutter 1989), to endow neural network models with controlled dynamic properties, in addition to their ability to deal with static patterns. With respect to the present neural network, we have not yet investigated how to optimize the storage capacity of the net (in terms of length and number of sequences) through the number of its hidden units (the granule cells), or how to control the length and form of the sequences through the choice of the values of the synapses of the Golgi-granule cell system.

Nevertheless, the model demonstrates that classical formal neural network modeling techniques, which are based on a few essential concepts, such as high connectivity in a net, weighted sums of activities, and non-linear units, can provide sufficient tools to construct models mimicking complex dynamic performances of biological systems when rules to achieve propagation delays also are included in them.

6.2 Biological standpoint

The present neural network model is based on the following general hypothesis: a movement or trajectory in real space is coded and memorized by the cerebellar cortex as a trajectory in the state-space of a neural network. We demonstrated that these performances can be supported by a neural network model that takes into account important biological constraints and therefore is biologically plausible. In particular, in the model presented here:

- the cerebellar cortex anatomy is closely reproduced by the network architecture (especially the presence of a very large number of granule cells making connection with a small number of Golgi cells),
- synaptic efficacies are either inhibitory or excitatory, according to their physiological nature,
- only parallel fiber-Purkinje cell synapses need be modifiable,
- learning is performed according to Hebbian rules rather than in form of a gradient-descent minimization.

The model includes an interpretation of the respective roles of the Golgi-granule cell system, and of the Purkinje cell layer. The Golgi-granule cell system, due to its feedback loop, generates a temporal sequence of patterns in response to an input pattern applied on the mossy fibers. The Purkinje cell layer, acting as a multiple perceptron, maps this sequence into a sequence of commands for the motor system. The input pattern presented on the mossy fibers is assumed to code both an intentional movement command coming from higher cerebral centers, and the context of the movement coming from sensorial sources. In the *retrieval phase*, characterized by the absence of activation of the climbing fibers, a temporal output sequence is generated on the Purkinje cell axons, which codes for the sequential commands necessary to perform or control the desired movement in the specified context. In the *learning phase*, when climbing fibers are activated, the parallel fiber-Purkinje cell synapses are modified under the control of the climbing fibers, in order to adjust the output sequence.

The neural network model makes use of propagation delays of nervous signals on parallel fibers. From a biological standpoint, the presence of significant propagation delays in these fibers is quite plausible. It is known (Eccles et al. 1967) that granule cell axons, which constitute these parallel fibers, have a small diameter (0.2–0.3 μm), and can extend over several millimeters. In such thin and long fibers, the propagation velocity of a nervous signal can be as small as a

few tens of centimers per second (see Gieger 1977), and thus can lead to propagation times of the order of tens of milliseconds. To generate sequences in the order of seconds that may be involved in the control of movement by the cerebellar cortex, would require the generation of sequences with a length of the order of 100 time steps. As shown in Sect. 3, this type of performance, under this length constraint, can be obtained by the sequence generation mechanism of the Golgi-granule cell system. Another argument for the importance of propagation delays has been suggested by Braitenberg (1990). He noted that the metric invariance observed in the organization of the cerebellar cortex network, as opposed to the merely topological invariance prevalent in other neuroanatomical structures, could reveal a crucial role of the times of signal transmission and reception in different points of the parallel-fiber beam.

In the model developed here, the mossy fiber input pattern **U** was assumed to be constant during iteration of the dynamics of the network. This is by no means a necessity, and the sequence generation ability, which is due to the delayed feedback, is preserved if input pattern **U** is changed during iteration. In such case, the same succession of patterns **U** need to be presented during retrieval and learning phases to have correct retrieval of a memorized sequence. In such operation, the evolution of **U** can be interpreted as monitoring the evolution of the context resulting from the movement that is being performed. A global feedback of this type could lead to greater stability or generalization capacity for the whole system: due to successive readjustments, a retrieved output sequence is prevented from gradually diverging from the corresponding memorized sequence, when iteration is started with a slightly corrupted input pattern **U**.

We saw that the length of the output sequences on the Purkinje cell axons can be controlled by presentation on the mossy fiber inputs of a complete deficity of activity, coded by a pattern consisting of only -1's. The system can realize this control of the length of the sequences in different other ways. For instance an adjustment of the fixed synapses of the Golgi-granule cell system could provide such a control. We mentioned that the number of granule cells and the values of the fixed synapses (and their amount of dilution) determine the average length of the sequences generated by the Golgi-granule cell system. This provides the Purkinje cell layer with sequences of a given average length, that would correspond to the quantity of patterns necessary to encode the information to perform or control a phase in a movement. A given sequence can also be terminated and another one initiated when the mossy fiber input pattern is simply changed during iteration of the dynamics. In such a schema, successive mossy fiber patterns applied to the system, could represent the coarse definition of a given movement. The role of the cerebellar cortex would be to complete this coarse coding, by inserting additional information in the form of sequences of patterns, in order to generate a fine implementation of the movement.

In the framework of the model, an interpretation for the divergence of the number of granule cells observed in the cerebellar cortex is provided. On the one hand, as output units of the Golgi-granule cell network, a large number of granule cells guarantees both large variety and good decorrelation for the sequences generated at this level. On the other hand, as input units of the Purkinje cell network, a large number of granule cells leads to the possibility of embedding a large number of patterns with almost unit probability of correctness, even with sign-constrained synapses. Phrased in formal neural network terminology, the role of the divergence of the granule cells is to implement the so-called thermodynamic limit, when the number of cells in a net becomes so high that any finite storage capacity can be achieved, even in the presence of constraints. This may not appear as an *optimal* way of devising a neural network, but it at least is an *efficient* one.

The model also offers elements to further interpret the cerebellar ability for coordination of movements. What has been considered here, in Sect. 1, is the set of granule cells that connects on a given Golgi cell. Such a Golgi unit can drive several Purkinje cells, each of which, in turn, supplies an output signal for the control of a given parameter in a movement. It is possible to consider several of these Golgi units, and introduce interactions among them through signals exchanged at the level of their Golgi cells. From a biological standpoint, such interaction between neighbouring Golgi units is plausible. This would make the sequences generated by a Golgi unit dependent on the behavior of the neighbouring units. In other words, sequences of patterns for the control of various movements, or various movement parameters, could be learned or retrieved in relation with one another. This type of mechanisms could provide a schema to represent and interpret performances of coordination of movements. We currently are investigating this feature of the model.

The proposed model stands for a global theoretical description of some general performances of the cerebellar cortex. Although many questions concerning the cerebellar function are not addressed by the present model, it nevertheless leads to several predictions that can be tested experimentally: propagation delays of neural signals on parallel fibers would have significant values, compared to Golgi cell response time; parallel fibers would tend to fire in a scattered and seemingly random fashion, rather than in coordinated beams; only parallel fiber-Purkinje cell synapses need be modifiable for learning of movement control abilities.

Acknowledgements. This work was supported by INSERM 88-9002 and DRET 88-1194 grants.

References

Albus JS (1971) A theory of cerebellar function. Math Biosci 45:247–293

Amit DJ, Wong KYM, Campbell C (1989) Perceptron learning with sign-constrained weights. J Phys A22:2039–2045

Braitenberg V (1990) Reading the structure of brains. Network 1:1–11

Chapeau-Blondeau F, Gouraud A, Chauvet G (1989) What type of neural network for the cerebellar cortex?; Proceedings of the Second International Conference on Neural Network and their Applications, Nîmes, France, pp 295–304 (in French)
Chauvet G (1986) Habituation rules for a theory of the cerebellar cortex. Biol Cybern 55:201–209
Chauvet G (1988) Learning abilities for a cerebellar Purkinje unit. Neural Networks 1 [Suppl 1]:244
Dehaene S, Changeux JP, Nadal JP (1987) Neural networks that learn temporal sequences by selection. Proc Natl Acad Sci USA 84:2727–2731
Dunin-Barkowski WL, Larionova NP (1985) Computer simulation of a cerebellar cortex compartment. Biol Cybern 51:399–406 and 407–415
Eccles JC, Ito M, Szentagothai J (1967) The cerebellum as a neuronal machine. Springer, Berlin Heidelberg New York
Fujita M (1982) Adaptive filter model of the cerebellum. Biol Cybern 45:195–206
Geiger SR (ed) (1977) Handbook of physiology. American Physiological Society, Bethesda, Md
Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA 79:2554–2558
Ito M (1984) The cerebellum and neural control. Raven Press, New York
Kuhn R, van Hemmen JL, Reidel U (1989) Complex temporal association in neural nets. In: Personnaz L, Dreyfus G (eds) Neural networks from models to applications. IDSET, Paris
Marr D (1969) A theory of cerebellar cortex. J Physiol 202:437–470
Massone L, Bizzi E (1989) A neural network model for limb trajectory formation. Biol Cybern 61:417–425
Minsky M, Papert S (1969) Perceptrons. MIT Press, Cambridge, Mass
Pearlmutter BA (1989) Learning state space trajectories in recurrent neural networks. Neural Comput 1:263–269
Pellionisz A, Llinas R (1982) Space-time representation in the brain. The cerebellum as a predictive space-time metric tensor. Neuroscience 7:2949–2970
Rumelhart DE, McClelland JL (1986) Parallel distributed processing, vol 1. MIT Press, Cambridge, Mass
Shinomoto S (1987) A cognitive and associative memory. Biol Cybern 57:197–206
Somopolinsky H, Kanter I (1986) Temporal association in asymmetric neural networks. Phys Rev Lett 57:2861–2864

Dr. F. Chapeau-Blondeau
Institut de Biologie Théorique de l'Ouest
Université d'Angers
10 rue André Boquel
F-49100 Angers
France