Cerebellar cortex: its simulation and the relevance of Marr's theory

TOBY TYRRELL AND DAVID WILLSHAW

Centre for Cognitive Science, University of Edinburgh, 2 Buccleuch Place, Edinburgh EH8 9LW, U.K.

CONTENTS

SUMMARY

Marr's theory of the cerebellar cortex as an associative learning device is one of the best examples of a theory that directly relates the function of a neural system to its neural structure. However, although he assigned a precise function to each of the identified cell types of the cerebellar cortex, many of the crucial aspects of the implementation of his theory remained unspecified. We attempted to resolve these difficulties by constructing a computer simulation which contained a direct representation of the 13 000 mossy fibres and the 200 000 granule cells associated with a single Purkinje cell of the cerebellar cortex, together with the supporting Golgi, basket and stellate cells. In this paper we present a detailed explanation of Marr's theory based upon an analogy between Marr's cerebellar model and an abstract model called the associative net. Although some of Marr's assumptions contravene neuroanatomical findings, we found that in general terms his conclusion that each Purkinje cell can learn to respond to a large number of different patterns of activity in the mossy fibres is substantially correct. However, we found that this system has a lower capacity and acts more stochastically than he envisaged. The biologically realistic simulated structure that we designed can be used to assess the computational capabilities of other network theories of the cerebellum.

1. INTRODUCTION

The cerebellum is a part of the brain that is thought to be involved in motor control. It has long been held, as a result of lesion studies, that it is implicated in the learning of motor coordination (Ito 1984; Gilman *et al.* 1981).

Although the gross function of the cerebellum is understood, there is no consensus on how it achieves this function. Its regular structure, which has under-

Phil. Trans. R. Soc. Lond. B (1992) **336,** 239–257
Printed in Great Britain

gone much detailed neuroanatomical investigation, provides many hints. There are four main schools of thought: the cerebellum either acts as a pattern associator (see, for example, Marr 1969; Albus 1971; Gilbert 1974), as a device for mapping between vectors (the tensor network theory of Pellionisz & Llinás (1982), described also in Churchland (1986)), as a biological timing device (Braitenberg & Onesto 1961), or as part of a circuit to implement classical conditioning (Thompson 1990).

Most theories of the cerebellum have been formulated at the algorithmic level; that is, mathematical equations are set up to simulate the action of the various cell types, without being directly subject to the constraints of the neuroanatomy. One such theory is due to the late David Marr (1969). He proposed that each output cell of the cerebellum controls an elemental movement of the body in response to the specific contexts in which the movement occurs; and that a process of associating these contexts with movement commands takes place. Although his proposed implementation was spelled out in much detail, it remains essentially a mathematical model and very few physical constraints were used. He did, however, suggest a neurobiological interpretation for the constituent elements of his theory, and he identified a particular type of synapse as constituting the modifiable element. The theory of Albus (1971) ascribes a similar role of pattern association to the cerebellum, but by means of a slightly different mechanism. However, to many people these two theories are identical, and they constitute the 'Marr–Albus' theory (Ito 1984) of the cerebellum.

Given that the cerebellar cortex has the most regular anatomy of any brain region and that there is a wealth of structural information available, we decided that it would be possible to construct an accurate full-scale model of a small part of the cerebellum that could be used to test the computational validity of theories of the cerebellum. We focused on Marr's theory because we felt it to be computationally the most tractable.

Our goal was to construct a computer model of the cells and connections influencing a single output cell that would embody as much of the anatomical structure as possible. Besides being useful as a tool to investigate the feasibility and performance of Marr's theory, the simulated structure would be also useful as a model of the cerebellum in its own right that could be adapted to test other theories.

The project involved a number of steps.

1. Collation and evaluation of current neuroanatomical data to establish the values of the parameters needed to construct the model.
2. Construction of a full-scale computer-simulated model of the basic cerebellar unit as identified by Marr. This unit comprises a population of some 13 000 mossy fibres that make synapses with 200 000 granule cells which contact one Purkinje (output) cell, together with a smaller number of supporting cells (Golgi cells, and basket and stellate cells).
3. Use of the simulated structure to test Marr's claims for his theory: that the cells of the cerebellum can interact in the way he outlined to efficiently associate input (mossy fibre) patterns with output (Purkinje cell) patterns, and particularly that each output cell can learn to respond to approximately 200 different contexts. Marr recognized that simulation would constitute the most direct method of testing, but in his day it was impossible to simulate a full-size system.

This paper makes three main contributions.

1. Marr's own paper is not easy to read, and we provide what we consider to be a clear, step-by-step explanation of it.
2. It describes how a computer simulation of part of the cerebellum can be constructed. This may seem a straightforward task, but in reality it is still difficult to obtain values for all the key anatomical parameters. It is also not always obvious how to generate, in the simulated model, the anatomical structure observed in real-life.
3. Using our model of the cerebellum, we were able to test Marr's theory by forcing his sometimes rather vague ideas to be integrated with the modelled neuroanatomy. In the process of implementing the theory, we identified several anomalies which we had to resolve, leading to changes in the basic model. Marr's estimate that each Purkinje cell can learn to respond to 200 different contexts is found to be of the right order of magnitude, even though some of the assumptions he used in obtaining that figure are shown to be incorrect.

The plan of the paper is as follows. In § 2, a brief survey of the anatomy of the cerebellum is given. Some associative memory theory is then described in § 3 and is used to explain Marr's ideas about the functioning of the cerebellum. In § 4 the steps taken to establish the parameters for the simulated model are given, together with a description of how it was constructed and the differences between it and Marr's original model. In § 5 we describe the simulation tests we carried out. The results are then discussed in § 6.

2. THE CEREBELLUM

(*a*) *The function of the cerebellum*

The cerebellum is a part of the brain that is involved in motor control. It is not essential to motor control, but it enables greater rapidity, smoothness, precision and complexity of movements. Animals and humans with damaged or destroyed cerebella are still able to perform movements, but these movements will be slow, inexact and uncoordinated (Gilman *et al.* 1981; Carlson 1977). The cerebellum seems to be responsible for the activation of large sets of sometimes spatially distinct muscles in a quick, well-timed and synchronized sequence. Humans are born with no capacity to perform many complex actions such as walking, writing and speaking, but can acquire the ability to perform them after extended practice. This process of acquisition, or learning, is thought to take

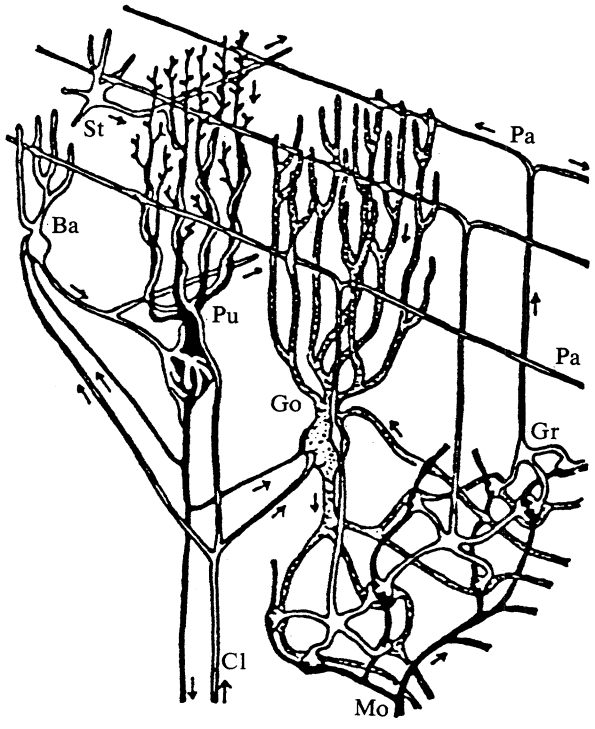

Figure 1. Cerebellar neurons. Pu = Purkinje cell, Go = Golgi cell, Gr = granule cell, Pa = parallel fibre, St = stellate cell, Ba = basket cell, Cl = climbing fibre and Mo = mossy fibre. From Llinás (1975). Copyright acknowledged to Scientific American Inc.

place in the cerebellum. The cerebellum is also involved with movements that are even more automatic such as the maintenance of stance and posture, and saccadic eye movements.

An example of the kind of movement in which the cerebellum is involved is given in Carlson (1977):

> For example, if you hold your arm straight out in front of you, it is possible for you to move it rapidly so that your hand describes a circle. Try this and move your arm as rapidly as you can. You will note that in doing so, you engage not only the muscles of your arm, shoulder and neck, but also those of your trunk and (especially if you stand) your legs. A phenomenal number of muscles are called into action, and at precisely the correct time. Just considering the arm movement alone, various muscles must begin and end their contractions at precisely the correct time in order to produce a smooth motion (after all a single muscle cannot produce a circular motion at the end of the arm).

In computational terms, the cerebellum can be seen as a device that relieves the cerebral cortex of the burden of conscious control of movements, freeing its computational capacity for other tasks. It also enables a more complex and coordinated control of the muscles than would be available with the cerebral cortex alone.

Finally, it needs to be appreciated that the vast majority of complex movements that we are able to perform have only been acquired after years of practice and experience, and are not hard-wired. During this learning there is a gradual transformation from total conscious cerebral control to an automatic unconscious execution of the movements involving the cerebellum.

(*b*) *The structure of the cerebellum*

There are two sets of inputs to the cerebellum, through the mossy fibres and through the climbing fibres. The mossy fibres are thought to relay information about the state of the body (positions of limbs, rotations of joints, resistances to rotations, etc.). The climbing fibres relay information from the inferior olive, and this information was thought by Marr to be the product of higher-level processing in the cerebral cortex. A third type of input to the cerebellum, through the aminergic fibres, has been discovered since Marr's time and so was not included in his model. This third type of input may signal reward, and could be incorporated fairly easily into Marr's model (Gilbert 1974) although it will not be considered here.

The inhibitory Purkinje cells are the only output cells of the cerebellum. Each Purkinje cell axon affects the contraction of an individual muscle, or group of muscles, in the body. The Purkinje cells are contacted directly by the climbing fibres. They are also contacted by the parallel fibres, the axons of the granule cells. The granule cells themselves are innervated by the mossy fibres, the second set of inputs to the cerebellum.

Besides the granule cells, the cerebellum has three other types of interneuron. These are the Golgi cells, the basket cells and the stellate cells.

An overview of the neuroanatomy follows, which should be read in conjunction with figures 1–7.

(i) *Purkinje cells*

Each Purkinje cell has a very flat and fan-like two-dimensional dendritic tree which intercepts and makes synapses with a large number (*ca.* 200 000) of parallel fibres (the axons of the granule cells (see figure 5)). It also receives synaptic contacts from several basket and stellate cells and a single climbing fibre.

(ii) *Climbing fibres*

Each Purkinje cell is innervated by just one climbing fibre which makes extensive contacts on the dendritic tree of that Purkinje cell (see figure 5). The contacts are sufficiently extensive that firing of the climbing fibre automatically induces firing of the Purkinje cell. A particular climbing fibre may innervate more than one Purkinje cell.

(iii) *Mossy Fibres*

The mossy fibres travel 'underneath' the cerebellar cortex proper, with each fibre sending out an occasional branch which 'ascends' to the cortex and then branches further to form a cluster of on average 7.5 axon terminals. Each of these axon terminals has contacts with, on average, 20 granule cell dendrites so that each mossy fibre cluster of axon terminals contacts of the order of 150 granule cells.

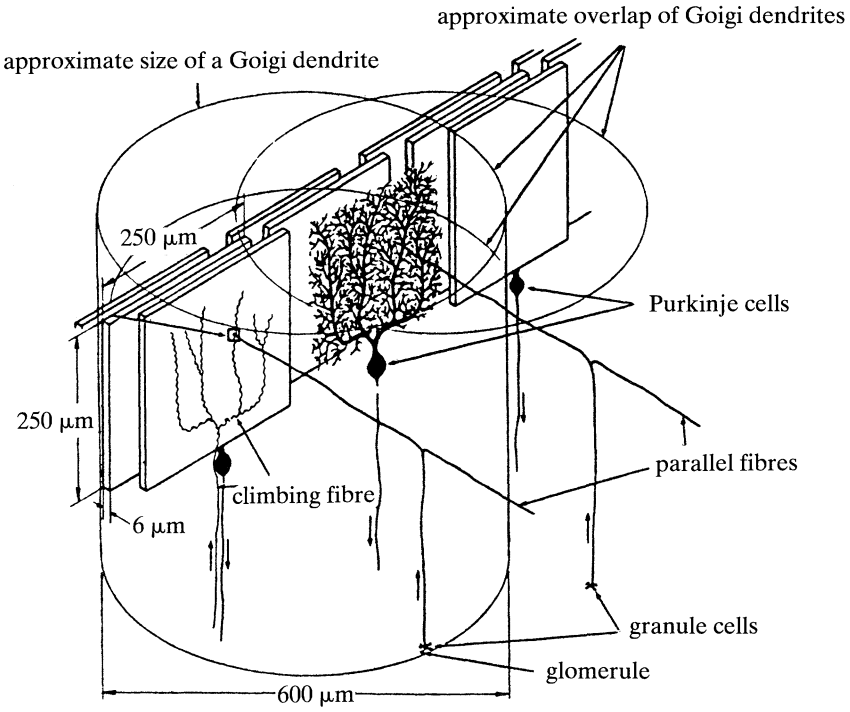

Figure 2. Cerebellar neurons. From Baron (1987). Copyright acknowledged to Lawrence Erlbaum Associates Ltd.

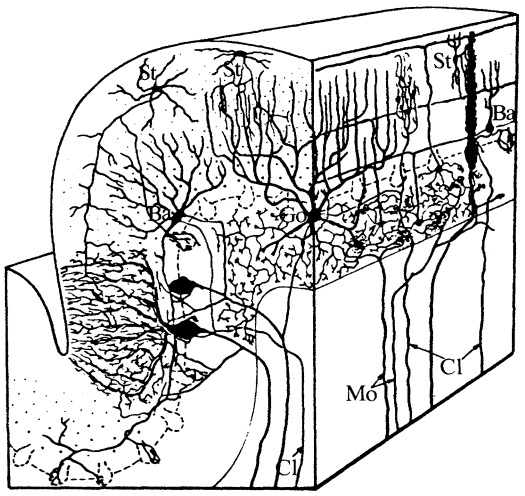

Figure 3. Cerebellar neurons. Mo = mossy fibre, Ba = basket cell, St = stellate cell, Cl = climbing fibre and Go = Golgi cell. From Eccles *et al.* (1967). Copyright acknowledged to Springer–Verlag.

(iv) *Granule cells*

The axons of the granule cells (the parallel fibres) are aligned perpendicularly to the flat dendritic trees of the Purkinje cells. This arrangement of parallel fibres and Purkinje cells provides the greatest possible number of parallel fibre to Purkinje cell contacts per unit volume, and also has the effect that very few parallel fibres contact an individual Purkinje cell more than once. Each granule cell is contacted by on average 4.5 mossy fibres.

(v) *Golgi cells*

They have two dendritic systems, one which ascends to take input from the parallel fibres and the other which descends to take input from the mossy fibres. Their axons branch profusely and make many inhibitory synaptic connections with granule cell dendrites (see figure 6).

(vi) *Basket and stellate cells*

These two types of neuron lie at different levels in the cerebellar cortex, and they make inhibitory connections with different parts of the Purkinje cell. Both basket and stellate cells are innervated by the same source (parallel fibres) and send their axons to the same destination (Purkinje cells). They are generally assumed to be functionally equivalent (see figure 7).

3. MARR'S THEORY OF THE CEREBELLUM

In his 'Theory of cerebellar cortex' Marr (1969) addressed the problem, described in § 1, of how the gross function of the cerebellum might be achieved by its neural machinery. This section presents an explanation of that theory (which is the first of the three contributions of our paper mentioned in § 1).

Marr suggested that the cerebellum learns the unconscious execution of movement through pattern association. The patterns being associated are those of proprioceptive information (state of the body) in the input axons, the mossy fibres, with those of motor control (activations of muscles) in the output neurons, the Purkinje cells. During learning, the conscious instructions as to which outputs to associate with the given mossy fibre context are carried along the climbing fibres to the appropriate Purkinje cells. After learning, the contexts alone will activate the relevant output patterns, and the execution of movements can be carried out automatically with no guidance from the cerebral cortex.

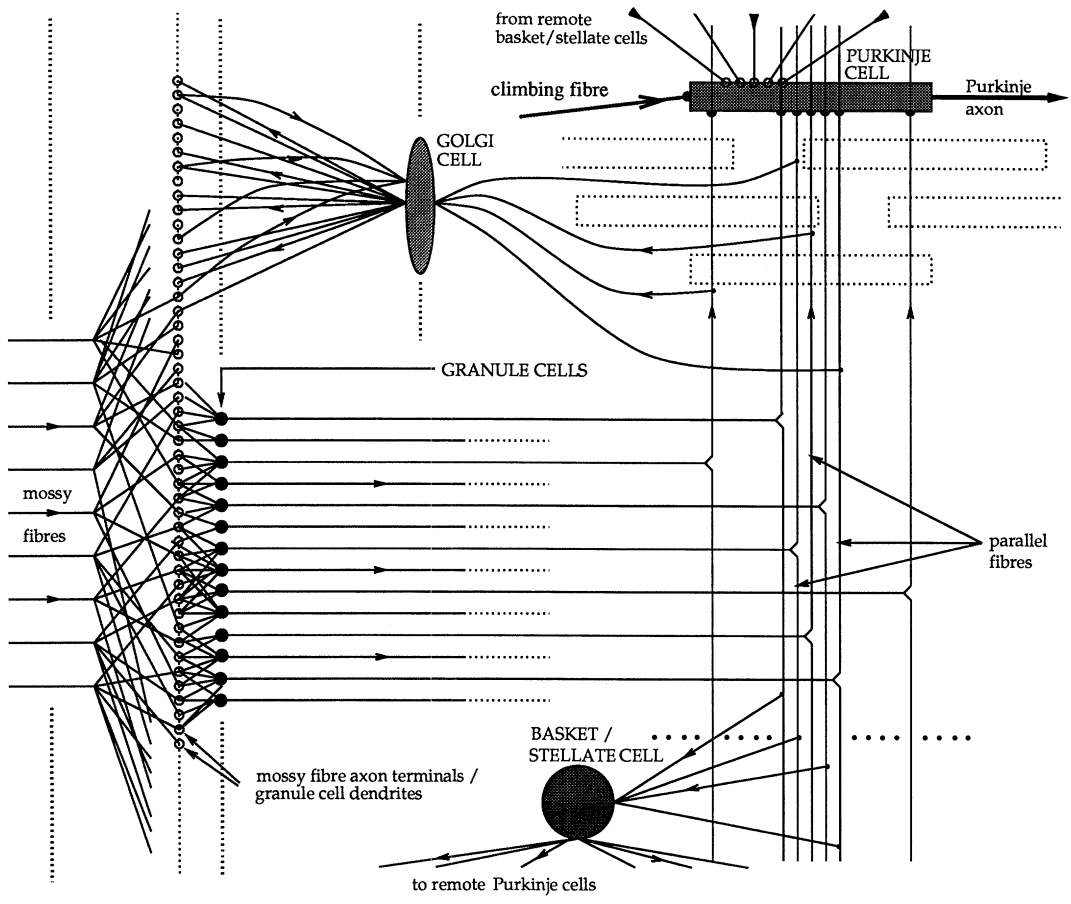

Figure 4. Schematic diagram of the arrangement of neurons in the cerebellum.

(*a*) *The associative net*

As a way of introducing Marr's theory, we describe a more abstract formalization of associative memory, the associative net (AN), also known variously as the associative matrix, correlation matrix, or Willshaw Net (Willshaw *et al.* 1969). The associative net is a simple computational device which acts as a pattern associator. Marr's whole theory can be viewed as an implementation of the associative net in the cerebellum (although he did not express it in these terms).

There are a set of input lines and a set of output lines with a set of binary-valued modifiable synapses at their intersections. Each input and each output line can be set to either a high (excited) or a low state.

The AN is able to form links or associations between patterns in its input and patterns in its output. On subsequent re-presentations of a stored input pattern, the net is able to make use of the associations stored to respond with the corresponding output pattern. The net therefore has two states: learning mode when it is forming associations, and discriminating mode when it is deciding whether or not to respond to an input pattern, and, if so, with what output pattern.

It is able to carry out this function in the following way.

1. Initial state. Each input line has a synapse with

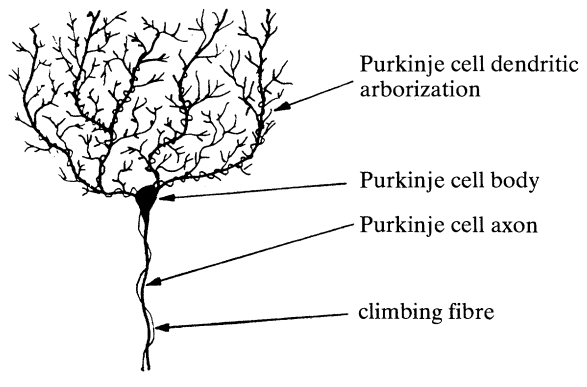

Figure 5. Intertwining of climbing fibre around extensive dendritic arborization of Purkinje cell.

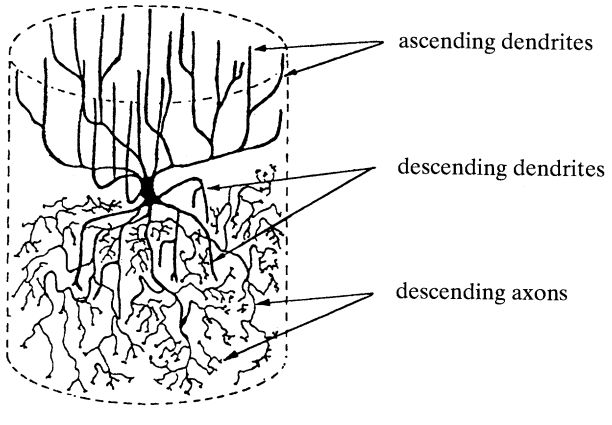

Figure 6. Structure of the Golgi cell. From Albus (1971). Copyright acknowledged to Elsevier Science Publishers Ltd.

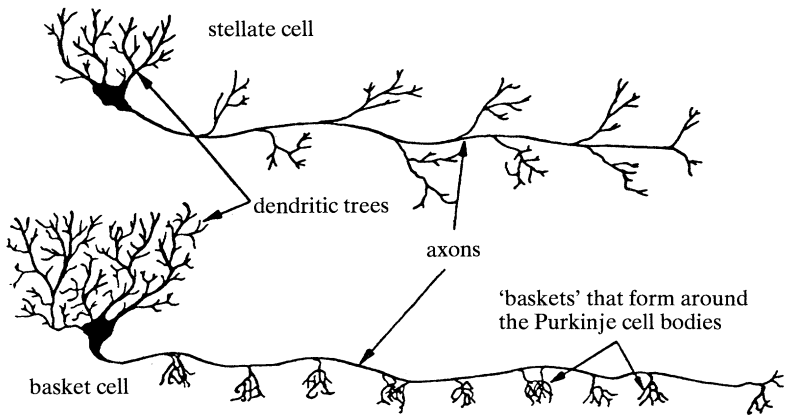

Figure 7. Structure of the basket and stellate cells.

each output line as shown. These synapses are all OFF initially (see figure 8*a*).

2. *Learning mode.* When a pattern on the input lines is presented together with the pattern on the output lines with which it is to be associated, all the synapses between excited input lines and excited output lines get switched ON (see figure 8*b*).

3. *Discrimination mode.* When the net is in discrimination mode, each output line has to decide whether or not it should be active in response to the given input pattern. It does this by summing the excitation coming through its previously modified synapses (thereby counting the number of modified synapses on activated input lines) and comparing that number with the sum of excitation in all the input lines (the number of activated input lines). If the two numbers are equal, then the output line is made active. When a learned pattern is later re-presented in the input lines, only those output lines that were part of the output pattern originally associated with it will be activated (see figure 8*c*). When an unlearned pattern is presented to the network, provided not too many

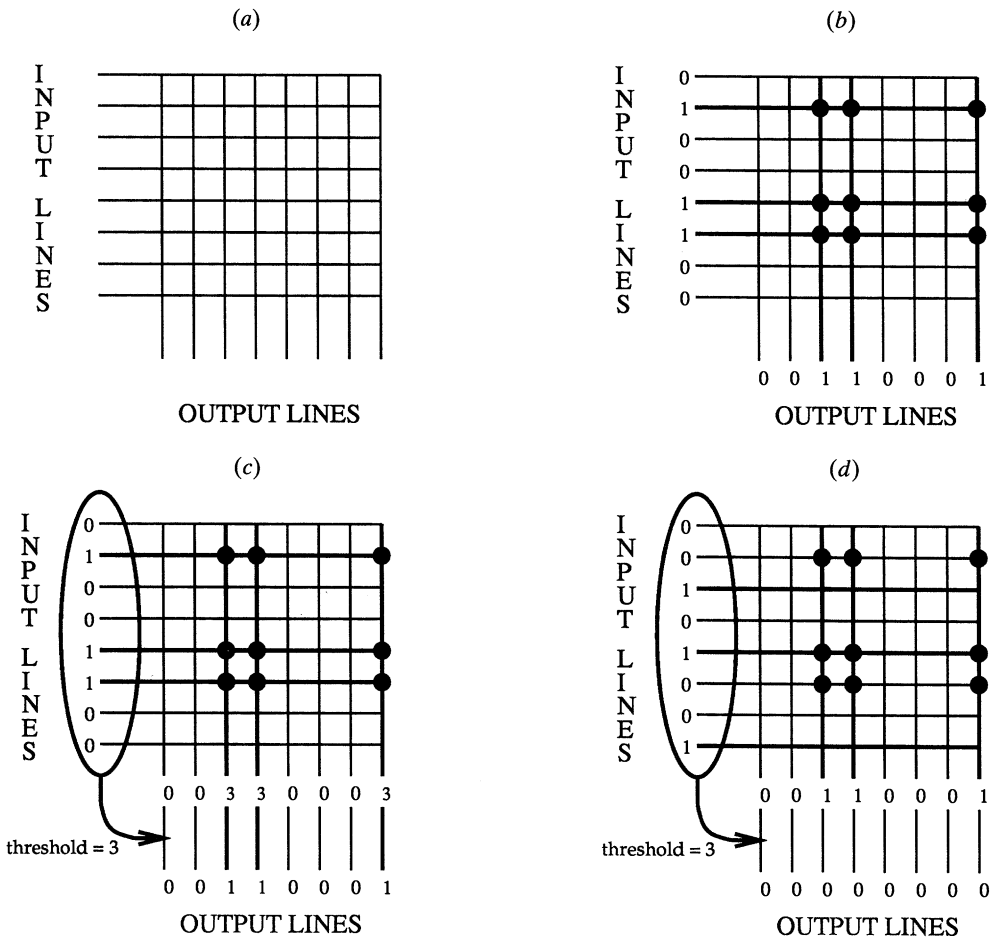

Figure 8. Operation of the associative net: (*a*) initial state; (*b*) learning to associate an input pattern with an output pattern; (*c*) subsequently presenting a learned pattern; (*d*) subsequently presenting an unlearned pattern.

synapses have been turned ON by other associations, no output lines will be active and so there will be no response from the system (see figure 8*d*).

(*b*) *The cerebellum as an associative net*

The following comparisons can be made between components of the associative net and the cells of the cerebellum: (i) the parallel fibres are the input lines of the AN; (ii) the Purkinje cells are the output lines; (iii) the synapses between parallel fibres and Purkinje cells are the synapses in the AN; (iv) the climbing fibres signal whether the net should be in learning or discriminating mode (i.e. they tell each output line (Purkinje cell) whether it should be active in the given input pattern); and (v) the basket and stellate cells perform the thresholding operation on the output lines, as explained below.

To implement the thresholding operation biologically, information about how many parallel fibres are activated and have modified synapses is required. This information is available to the output cell directly, through the depolarizing effect that they can be assumed to have on the Purkinje cell dendrites. However, the total number of activated parallel fibres, which is also required, is unknown. Marr assumes that this information is provided by the basket and stellate cells which sample the parallel fibre activity and provide an inhibitory signal which is proportional to the total excitation in the parallel fibres. There will be a competition between the excitation received through the parallel fibres with modified synapses and the inhibition received through the basket and stellate cells. This competition will result in the Purkinje cell (output line) firing only when it is part of an output pattern associated with the input pattern. The implementation of this is described in § 4.

(*c*) *The improved associative net*

The analogy with the associative net, as developed so far, is not sufficient to explain the whole anatomy of the cerebellar cortex (as described in § 2*b*). The theory still needs to account for the existence of the granule cells (why do the mossy fibres not synapse directly onto the Purkinje cells, instead of indirectly via the granule cells?), the complex nature of the connections between the mossy fibres and the granule cells, and the existence of the Golgi cells. As will be explained below, these provide the machinery for solving three specific problems in the associative net scheme.

1. Saturation. The major problem is that the capacity is severely limited. As more and more associations are made, more and more of the synapses become modified. As the proportion of modified synapses increases then the probability of the net making incorrect responses to unlearned input patterns (i.e. producing false positives) also increases. This phenomenon is known as saturation (see figure 9). Saturation imposes a limit on the capacity of the network. As more associations are learned, the performance of the system gradually degrades.

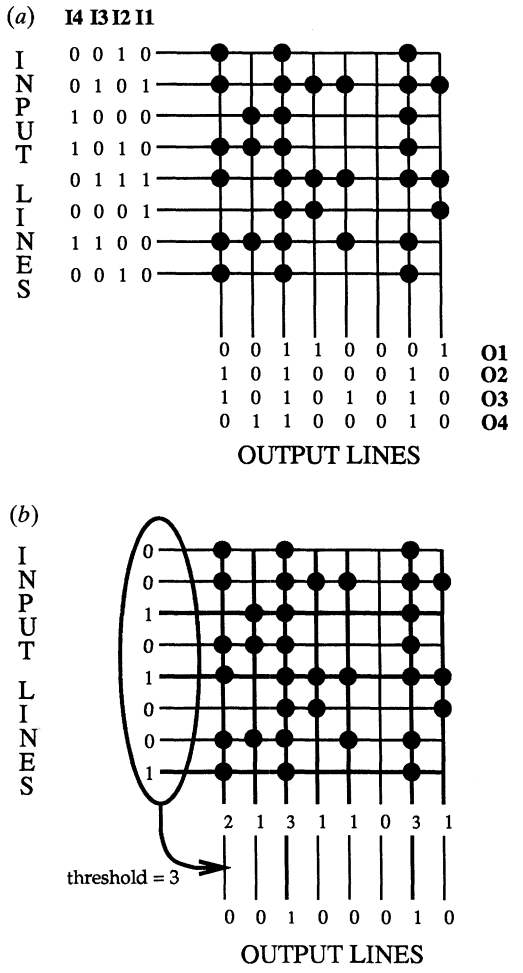

Figure 9. Saturation in the associative net: (*a*) turning on more synapses with more associations; (*b*) two false positives when presenting an unlearned pattern.

2. Subsets. A problem with the associative net as applied to this case is that it responds to subsets of learned contexts as well as to the learned contexts themselves (see figure 10). Although in many other cases this phenomenon is a desirable property of associative memories, it is our feeling that for this task it is better to construct the net so that it only responds to a subset if that subset has been learned explicitly. For example, if the cerebellum had learned an association between mossy fibres signalling the three states 'mouth open', 'hand holding cup', and 'hand near mouth' and a Purkinje cell contributing to the action 'turn hand to pour contents of cup into mouth' then it would be undesirable to trigger the response for a subset of the mossy fibre pattern (e.g. just 'mouth open' and 'hand holding cup' alone).

3. Separating similar patterns with biological thresholding. The anatomy is not so precise and exact that each basket or stellate cell makes one, and only one, contact with each parallel fibre, but rather each cell has sparsely distributed sets of dendrites which can only sample the activity in the parallel fibres (see figure 7). These cells therefore cannot provide an exact measure of the input activity. To make certain that a Purkinje cell responds to all learned patterns it is necessary to reduce the threshold on the number of modified

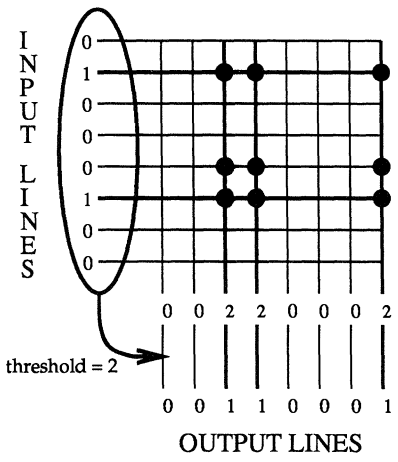

Figure 10. Responding to a subset in the associative net.

synapses on activated input lines that must be exceeded for a Purkinje cell to fire. This produces a problem of confusing similar input contexts. Since slightly lower thresholds are required to accept all learned contexts, it will become more likely that some unlearned contexts which are very similar to learned contexts will be recognized as well.

How can these problems be overcome; or at least their effects be reduced? To solve the first problem, the proportion of the input neurons that are excited per pattern needs to be reduced. To achieve this reduction it is obviously not sensible to just turn off the activity in some of the input lines since then different patterns would become identical. A better idea is to transform the input into a much larger set of neurons. A similar number of neurons in the larger set can be excited (preserving information and therefore the ability to discriminate between patterns), while at the same time a smaller proportion of the input neurons will be excited in the larger set (thus turning on a smaller proportion of the synapses with each association and thereby increasing the capacity of the net). This scheme is shown in figure 11, with the method for deciding the mapping between excited first layer neurons and excited second layer neurons left at present as a black box.

This transformation of inputs into a much larger set will help with the second and third problems if it has

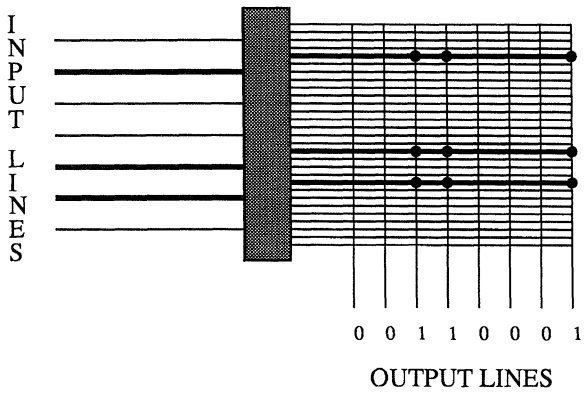

Figure 11. Expanding the input of the associative net into a larger size pattern, but with a similar number of excited neurons.

the following properties: (i) given that pattern A in the first input layer produces pattern B in the second layer, a subset of A must not produce a subset of B; (ii) two similar first layer input patterns must be transformed into patterns in the second layer that are less similar.

The transformation therefore needs to be complex. Just repeating the first layer pattern several times over in the larger second layer pattern and then inhibiting some lines would not help greatly with subset recognition and pattern separation.

(*d*) *The cerebellum as an improved associative net*

There is a need for an expansion of sufficient complexity of the original input pattern into a pattern in a larger set of input lines with a lower average excitation. Here this need is linked to the existence of the granule cells, the complicated mossy fibre to granule cell connections and the Golgi cells.

There are approximately 13 000 mossy fibres that contact the approximately 200 000 granule cells that then synapse with one particular Purkinje cell (see § 4*a*). Thus there are about 15 times more granule cells than mossy fibres for a particular Purkinje cell. The mossy fibre to granule cell connections are shown schematically in figure 4 and were discussed in § 2*b*. Each granule cells is contacted by on average 4.5 mossy fibres.

Marr assumed that each mossy fibre to granule cell synapse is unmodifiable and has unitary weight. Since each granule cell is contacted by more than one mossy fibre, the fraction of granule cells receiving some activation (α_g) will exceed the fraction of activated mossy fibres (α_m). Marr proposed that the fraction α_g of active parallel fibres can be made less than α_m by means of the inhibition supplied by the Golgi cells. By providing inhibition in proportion to the level of granule cell activity that would result without inhibition, they are able to transform mossy fibre patterns of widely different levels of activity into granule cell patterns with roughly the same level. Marr suggested that: (i) the Golgi cell descending dendrites, which are contacted by the mossy fibres, provide fast prediction of rapidly changing granule cell activity (the uninhibited granule cell activity would be proportional to the mossy fibre activity), and (ii) the Golgi cell ascending dendrites, which are contacted by the parallel fibres, provide more accurate estimates of the actual granule cell activity for fine-tuning of the inhibition when the mossy fibre input is more stable. The assumed effect of Golgi cell inhibition is shown in figure 12.

Because each granule cell is excited by 1–7 mossy fibres (i.e. each granule cell samples a subset or, as Marr called it, a codon, of the mossy fibres) and is also inhibited by Golgi cells, the transformation between first and second layers is sufficiently complex to allow for rejection of subsets and pattern separation (as demonstrated in § 5). Changing of a few mossy fibre inputs will affect the excitation of many granule cells, and for many of those it will make the difference as to whether or not they survive the Golgi cell inhibition.

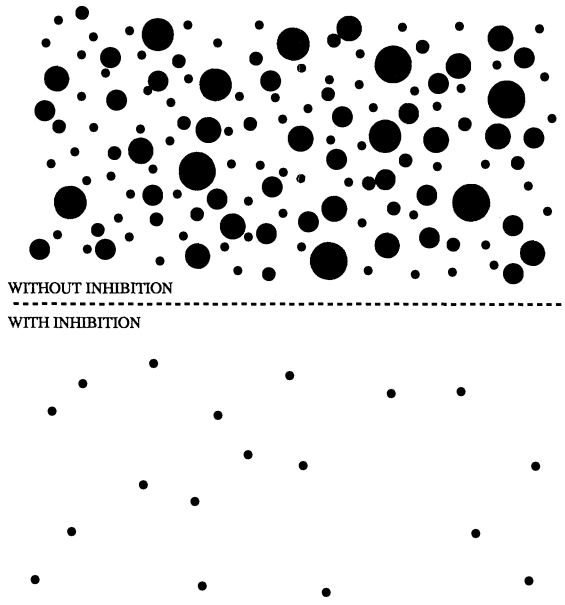

Figure 12. Assumed activity in the granule cells with and without Golgi cell inhibition. Larger circles correspond to higher excitation.

1. Saturation. To lessen the effect of saturation, the fraction α_g of granule cells active in any input pattern must be less than the fraction α_m for the mossy fibres.

2. Preservation of information. To ensure that the number of possible granule cell patterns be no less than the number of possible mossy fibre patterns, α_g must be large enough to satisfy

$$\alpha_g(\ln \alpha_g) \geqslant (N_m/N_g)\alpha_m(\ln \alpha_m),$$

where N_m, N_g are the numbers of mossy fibres and granule cells innervating one Purkinje cell.

3. Pattern separation. The parameter that is crucial in determining whether an output cell should fire is the ratio of the number of activated input lines onto modified synapses to the total number of activated input lines. Therefore an appropriate measure of the separation between two patterns is the number of fibres at which activity differs divided by the mean number of active fibres, called θ here. For two patterns of length N in which the probability of a component having value 1 is α, θ equals 2 $(1-\alpha)$, a decreasing function of α. Therefore, to obtain increased pattern separation in the granule cell layer as compared to the mossy fibre layer,

$$\alpha_g < \alpha_m.$$

Note that this constraint is equivalent to the condition for reducing the effect of saturation.

These two constraints provide upper and lower bounds on the value of α_g.

To summarize, the action of the mossy fibre to granule cell transformation has certain computational advantages, which are bought at the cost of the more elaborate machinery required. In terms of the three problems described in § 3*c*, these are: (i) saturation: having a lower level of activity in the transformed input patterns allows more associations to be stored reliably; (ii) subsets: because all transformed patterns will have roughly the same level of activity (regardless of α_m) and the transformation is complex, the pattern into which a mossy fibre pattern A will be transformed will be completely different from that into which a subset of A is transformed; (iii) pattern separation: because the transformation is complex, the granule cell patterns will be more separated than the mossy fibre patterns provided that α_g is less than α_m.

4. CONSTRUCTION OF THE SIMULATION

This section presents the second of the three contributions of this paper mentioned in § 1, i.e. how we constructed an anatomically realistic simulation of part of the cerebellar cortex.

Owing to the large number of cells in the cerebellum, we were restricted to modelling only a small section of it: that which contains all the cells forming the two pathways (mossy fibre and climbing fibre) to a single output (Purkinje) neuron. Showing that each output neuron can learn when to fire and when not to fire is sufficient to demonstrate that the network as a whole can learn to produce the correct output patterns for the respective input patterns; that is, it can learn to associate patterns and recall the associations correctly. This follows the approach of Marr, who assumed that the basic unit of the cerebellum is a Purkinje cell together with all the cells contacting it. We give here details of how the model was constructed, and the results we obtained with it are discussed in § 5.

The following items are required for a complete specification of the model (Lippmann, 1987): (i) the net topology (i.e. the numbers of cells and the connections between them); (ii) the node characteristics (i.e. the function of its input that each cell uses to determine its response); and (iii) the training or learning rules (i.e. the rules governing changes in the synaptic weights).

(*a*) *Net topology*

Most of the work in constructing the simulation involved producing the correct connectivity between the cells. This is a complex process, and only a short description of how it was done can be given here.

Although the cerebellum has a fairly regular structure compared with other parts of the brain, there is considerable variation in the distribution of dendrites, numbers of dendrites, lengths of axons, numbers of connections, etc. among cells of the same type. Nothing is known rigidly or exactly. This lack of certainty was represented in the model by incorporating randomness into the numbers of dendrites, the positions of dendrites around the cell body, and so on.

Owing to the large number of parameters in the simulation whose values are only sketchily known, exploration of all regions of the parameter space would have been impossible and so we decided to explore only the most favourable part of parameter space. We looked at the different estimates of the anatomical measurements that have been proposed and chose those which were most likely to allow the

simulation to work. It should be stressed that much of the anatomical information we used is sketchy, and that in many cases we had to make what we felt were plausible assumptions. Because of these factors, the simulation could only give us results of the nature, 'what we know of the anatomy of the cerebellum is not incompatible with the theory that . . .', rather than of the nature, 'the anatomy of the cerebellum can be proven to implement . . .'.

Although the schematic diagram of cerebellar structure in figure 4 attempts to portray the cerebellum in two dimensions, in reality the granule cell bodies, Golgi cell bodies, etc. are not arranged in lines, or even in planes, but rather are positioned in three-dimensional layers. One simplification we made was that when generating cells and forming connections between them, we replaced the layers that lie in three dimensions by planes in two dimensions (figure 13).

As far as possible, we represented the function of each type of cell as a computation carried out by individual cells at the cellular level, rather than treating all the cells of one type as a whole. The exception was the sampling by Golgi cells and by basket or stellate cells 'outside' the model. The processes from cells in both these classes extend a long way outside the modelled area of the cerebellar cortex. To model these connections explicitly would have meant increasing the size of the simulation by a factor of 20 or so (i.e. to about 4×10^6 granule cells). To make the simulation tractable we modelled the effect of the 'external' mossy fibres and granule cells sampled by the Golgi and basket or stellate cells implicitly. This introduced an element of arbitrariness into the simulation, in that we had to estimate how different the external and internal excitations would be, but this was unavoidable. We decided to use a maximum of 5% diference between the average excitations of internal and external cells.

The model was built up by generating granule cells in a layer beneath the Purkinje cell, giving their axons (the parallel fibres) random lengths, and then retaining only those which were long enough to reach to the Purkinje cell. The granule cells that remained were each given a random number of dendritic terminations positioned randomly around the cell body. A layer of mossy fibres (all those that might contact the relevant granule cells) was then generated. Each mossy fibre was given a random number of axon terminals arranged around the cluster centre, each terminal being a random distance away from the centre. The mossy fibre to granule cell connections were then formed by linking each granule cell dendritic termination to the closest mossy fibre axon terminal. Finally the Golgi cell bodies were randomly positioned in a separate layer. They were given two sets of dendritic terminations and one set of axon terminals positioned randomly around the cell bodies. These were then connected to the closest parallel fibres, mossy fibre terminals and granule cell dendrites respectively.

(i) *Purkinje cell*

The key parameters are the shape and dimensions of the dendritic tree, about which there is little disagreement (Eccles *et al.* 1967; Ito 1984). The dendritic tree is a flat structure with a width of 250 μm (Eccles *et al.* 1967) which is oriented perpendicularly to the large number of parallel fibres crossing it.

(ii) *Granule cells*

The key parameters are: (i) the length of the parallel fibres traversing the Purkinje cells; (ii) the number of granule cells innervating one Purkinje cell; and (iii) the number and dimensions of the granule cell dendrites.

One figure for the number of granule cells innervating one Purkinje cell is 200 000 (Eccles *et al.* 1967). Fox *et al.* (1967) give a similar figure of 120 000. Some estimates of parallel fibre lengths give a range from 2000 μm to 3000 μm (Eccles *et al.* 1967). Albus (1971) estimates ≈ 3000 μm. To obtain a plausible distribution of granule cells, we carried out the following procedure.

We calculated the average spacing between granule cells, which was that which would result if 200 000 granule cells were distributed regularly within a 2500 μm × 250 μm rectangle (defined by the average length of a parallel fibre and the extent of the Purkinje cell dendritic tree assumed). This gave a value of 1.77 μm for the spacing.

To cater for the observed variation in parallel fibre length we then arranged granule cells at this spacing within a rectangle of 3000 μm × 250 μm (defined by

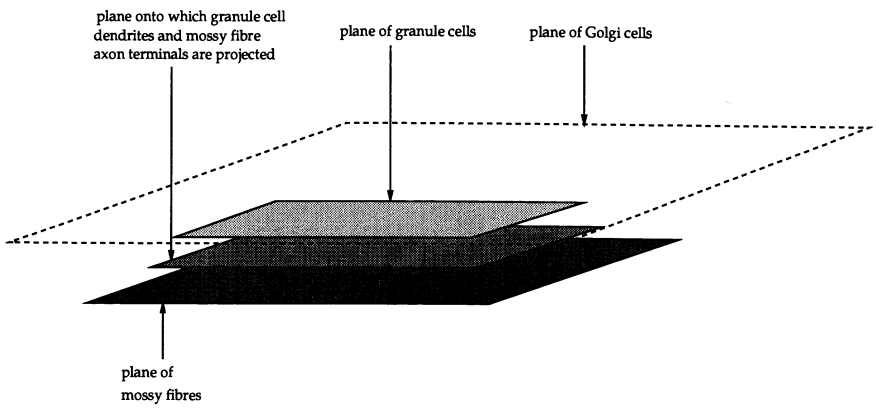

Figure 13. Assuming planes rather than layers in order to form connections easily.

the maximum length of a parallel fibre and the extent of the Purkinje cell dendritic tree assumed).

Each of these granule cells was then given a parallel fibre with a length randomly chosen in the range 2000–3000 μm, and those granule cells with a parallel fibre that did not reach the Purkinje cell were discarded.

By this procedure, approximately 200 000 granule cells reach the Purkinje cell with a spread of parallel fibre lengths. Eccles *et al.* (1967) and Albus (1971) assumed that almost all of these parallel fibres synapse with the Purkinje cell. We made synaptic contacts between every parallel fibre arriving at the Purkinje cell.

As well as putting out an axon, each granule cell puts out a number of short dendrites, claws, that contact the mossy fibres in a plane. There is some disagreement about the number and the length of the claws. We assume here that both quantities are randomly distributed. Following Marr, the number of claws is assumed to be between one and seven, with an average of 4.5 and the distance from the cell body to each claw is randomly chosen in the range 0–30 μm, and in a random direction.

(iii) *Mossy fibres*

The key parameters are: (i) the number of mossy fibres; and (ii) the number and distribution of mossy fibre axonal terminals in each cluster.

Marr uses various observations about divergence and convergence at the mossy fibre to granule cell interchange to estimate that approximately 6000 mossy fibres affect one Purkinje cell. He guesses that this number will be increased to 7000 by edge effects. We accounted for edge effects by assuming that these 6000 mossy fibres are found within the 2500 μm × 250 μm rectangle, which gives a cell separation of 10.2 μm assuming uniform spacing. The mossy fibre terminals occur in clumps of some 20 rosettes per mossy fibre, each on a stalk (Eccles *et al.* 1967; Fox *et al.* 1967). Given that each axon terminal from one mossy fibre is on a stalk of length 0–120 μm, and the claw length is 0–30 μm, the area of space containing mossy fibres that might influence the granule cell population defined above (which were assumed to originate from an enlarged rectangle) then becomes enlarged by twice the stalk length and twice the claw length to a rectangle of dimensions 3300 μm × 550 μm. The method of making contacts is now to place both granule cell claws and mossy fibre axon terminals in the same notional plane and join each claw to the closest axon terminal. Mossy fibres that make no contacts are then discarded. The effect of these two factors (increasing due to edge effects and discarding those making no contacts) is to increase the size of the mossy fibre population from 6000 to 13 000.

(iv) *The basket and stellate cells*

The cells assumed by Marr to set Purkinje cell thresholds are the off-beam cells (Eccles *et al.* 1967), that sample activity remotely. Because of the long-ranging dendrites of these off-beam basket and stellate cells, they sample cells that are too distant for them to be modelled explicitly. Given that P_I is the average activity of the parallel fibres internal to the model, the external activity P_E is calculated as $P_E = P_I \times (0.95 + \varepsilon)$, where ε is the average of two random numbers in the range 0.0–0.10.

(v) *Golgi cells*

Marr's specification of the parameters for the Golgi cells is the most vague and disputable part of his theory.

The key parameters are: (i) the number and distribution of the Golgi cells and their morphology; (ii) the number and distribution of the contacts made by the ascending and descending dendritic trees; and (iii) the number and distribution of the contacts made by the axonal system.

(vi) *The number and distribution of the Golgi cells and their morphology*

Marr took his interpretation of the Golgi cell's structure from Eccles *et al.* (1967), and so assumed that the Golgi cells partition up the cerebellar cortex (in the plane of the parallel fibres) into small, contiguous, non-overlapping, tessellated, hexagonal prisms (see figure 14). However, this account is not only biologically implausible but inconsistent. Eccles *et al.* (1967) took the diameter of the non-overlapping hexagons to be 700 μm, but to be compatible with the widely accepted figure of one Golgi cell per nine to ten Purkinje cells, this figure would have to be nearer 200 μm.

In his related theory, Albus (1971) assumed Golgi cells with roughly circular dendritic trees, with an average overlap of nine Golgi cells at any point on the cortex, and a diameter of ≈600 μm (see figure 15). As well as being more biologically plausible, the figures used are consistent with one Golgi cell per nine to ten Purkinje cells.

We calculated the diameter of the Golgi cell dendritic tree as follows. Looking from above, the average area of cortex occupied by a Purkinje cell is

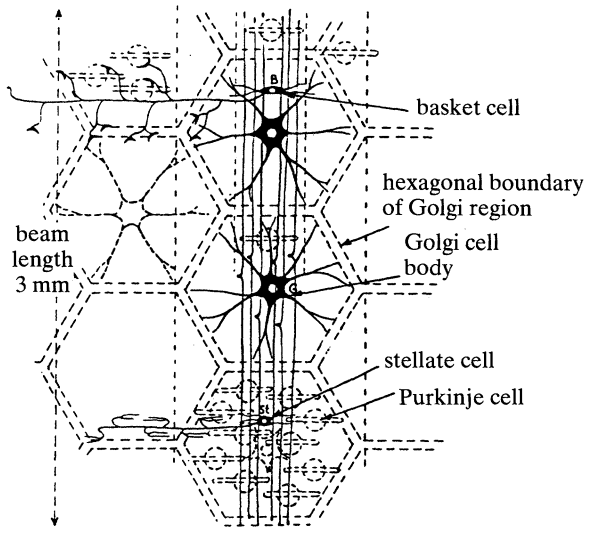

Figure 14. Marr's interpretation of the arrangement of Golgi cells. From Eccles *et al.* (1967). Copyright acknowledged to Springer-Verlag.

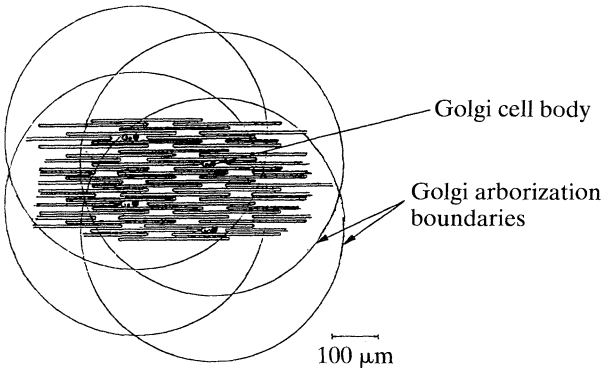

Figure 15. Albus's interpretation of the arrangement of Golgi cells. From Albus (1967). Copyright acknowledged to Elsevier Science Publishers Ltd.

$(250 + 50)\ \mu m \times 9\ \mu m$, or $2700\ \mu m^2$, where 9 μm is the perpendicular distance from one Purkinje cell dendritic sheet to another, and 50 μm is the spacing between the ends of two Purkinje dendritic spreads (Eccles *et al.* 1967). One Golgi cell per ten Purkinje cells thus requires one Golgi cell per $27\,000\ \mu m^2$, and so, assuming that on average nine Golgi cells overlap each point in the cortex, the area covered by each Golgi cell is $243\,000\ \mu m^2$. This gives a diameter of approximately 550 μm. The positions of the Golgi cells were generated by creating grid points 165 μm apart (so as to provide an average of one Golgi cell per nine to ten Purkinje cells) and then displacing the cells from the grid positions by a random amount of up to 50 μm. The dendritic and axonal terminations were then each placed at a random direction from the cell body of between 0 and 275 μm (corresponding to a diameter of 550 μm).

(vii) *Contacts on the descending dendritic tree*

Marr assumes that each Golgi cell has a 10% chance of sampling each mossy fibre beneath it with at least one of its dendrites; i.e. each mossy fibre has a 90% chance of not being sampled by this cell. We calculated the total number of mossy fibre axon terminals within a circle of radius 275 μm to be ≈17 000. The average number of axon terminals per cluster is taken as 3.75†.

The number *gdd* of Golgi cell descending dendrites needed so that each mossy fibre has a 10% chance of getting sampled at least once is then given by the equation

$$0.90 = \left\langle 1 - \frac{3.75}{17\,000} \right\rangle^{gdd}. \tag{1}$$

Therefore

$$gdd \approx -\frac{17\,000}{3.75} \ln(0.90) \approx 500. \tag{2}$$

This figure is of the same order of magnitude as the figure of 200 descending dendrites quoted by Pellionisz & Szentagothai (Ito 1984). In the simulation, each Golgi cell was given between 400 and 600 descending dendritic terminations at distances randomly distributed between 0 and 275 μm away from the cell body. These were then connected to the nearest mossy fibre axon terminal.

(viii) *Contacts made by the Golgi cell axon*

Albus assumes Golgi cells with 600 μm diameter axonal arborizations that inhibit 100 000 granule cells. Marr, assuming smaller size axonal arborizations, assumed that all the 4500 granule cells beneath a Golgi cell are inhibited.

With a probabilistic distribution of axons, a very large number of axons would be required for a high probability of inhibition of all the granule cells. We assumed the following.

1. The Golgi cell axon terminations are distributed within a circle of radius equal to that of their dendritic trees. There are 17 000 mossy fibre axon terminals (where the granule cell dendrites are contacted) within that area.
2. Each Golgi cell axon termination makes contact with one mossy fibre axon terminal, randomly chosen within the extent of the Golgi axonal tree. The Golgi cell terminal inhibits all the granule cell dendrites which the mossy fibre terminal contacts.
3. Eighty per cent of the granule cells beneath a particular Golgi cell are inhibited, via at least one of their four dendrites‡, by that Golgi cell.

The minimum value *ga* for the number of Golgi cell axons is then given by the equation

$$0.20 = \left\langle 1 - \frac{4.0}{17\,000} \right\rangle^{ga}. \tag{3}$$

This leads to

$$ga \approx -\frac{17\,000}{4.0} \ln(0.2) \approx 7000. \tag{4}$$

In the simulation each Golgi cell was therefore constructed with between 6000 and 8000 axons, distributed randomly within 275 μm of the cell body. Pellionisz & Szentagothai (Ito 1984) estimated 4800 axons, which is of the same order as our figure.

(ix) *Contacts on the ascending dendritic tree*

This tree contacts the parallel fibres. Since the parallel fibres are much longer than the diameter of the Golgi dendritic tree, its width rather than its area is relevant. If 200 000 parallel fibres intersect the dendritic tree of a Purkinje cell which is 250 μm wide, then 440 000 will intersect the ascending Golgi dendritic tree, which is 550 μm wide. The only data to suggest the number of ascending dendrites come from the various observations that there are more ascending then descending dendrites. We assumed, in line with Marr, that each Golgi cell samples approximately 10% of the parallel fibres passing through the dendri-

† To allow for the fact that the mossy fibre clusters are quite wide, it is assumed that on average only half of the axon terminals of each cluster will fall in the range of any particular Golgi cell.

‡ Although there are on average 4.5 dendrites per granule cell, they are not very wide ranging and here we have assumed that an average of about four dendrites per granule cell will fall inside the Golgi cell area.

tic tree, and each Golgi cell was therefore given a random number of ascending dendrites, chosen from the range 35 000–53 000. The parallel fibres that were contacted were chosen at random.

This concludes the description of how the static structure was generated in the computer simulation. The rest of this section describes how the structure is used.

(*b*) *Node characteristics*

All the synapses, except the modifiable ones between the parallel fibres and the Purkinje cell, are assumed to have unitary weights.

(i) *Golgi cells*

These are assumed by Marr to make two estimates of the average excitation received by the granule cells from the mossy fibres, which can be thought of as the number of cells that would fire in the absence of inhibition. Marr assumed that the higher of these two estimates is taken to determine the amount of inhibition to be applied. We had to modify extensively his suggestions as to how such a scheme might work. He had assumed one-to-one Golgi to granule cell connections within the smaller, hexagonal, non-overlapping Golgi compartments, which is biologically unrealistic. Our proposal is as follows: the estimate obtained by the ascending system of Golgi cells is obtained directly by measuring the proportion of the parallel fibres that this cell contacts that are active. The estimate supplied by the descending system is obtained by measuring the proportion of the mossy fibres contacted that are active, and this estimate is then multiplied by a constant factor representing the mean number of dendrites per granule cell. Both values include an influence from 'external' sampling which could be up to 5% different from the internal estimate.

The amount of inhibition supplied by a single Golgi cell is then calculated as

$$I = f_1 E + f_2, \tag{5}$$

where E is the larger of these two estimates of parallel fibre activity (in line with Marr's suggestion) and f_1 and f_2 are two constants whose values are obtained empirically so as to give a good mapping between the mossy fibre excitation values and the final (inhibited) granule cell excitation values.

(ii) *Criteria for a good mapping*

It is desirable to keep the average excitation of the granule cells after inhibition as low as possible so as to maximize the capacity of the net. Marr suggested that the target excitation should be about 1%. However, the more a granule cell pattern is reduced by inhibition, the less information from the original mossy pattern carries through and the more likely it is that two initially different granule cell patterns will get inhibited down to the same post-inhibitory granule cell pattern and so be impossible to discriminate between. Also, a desirable mapping will map mossy fibre patterns with higher than average excitation onto granule cell patterns with a similar higher than average excitation so as to minimize the information lost; similarly for patterns with lower than average excitation.

After some tests, we decided on an average target excitation of just over 1.0%. This figure gave the optimum balance between the factors of needing low excitations so as to increase the capacity and needing high excitations so as to avoid confusion of patterns.

(iii) *Granule cells*

These sum their excitatory inputs (from the mossy fibres) and inhibitory inputs (from the Golgi cells), and fire if the result is greater than zero.

(iv) *Purkinje cells*

These sum their excitatory inputs (from the parallel fibres) and their inhibitory inputs (from the basket and stellate cells), and fire if the result is greater than zero.

(v) *Basket and stellate cells*

These are presumed to sum the excitation coming through their dendritic connections with parallel fibres and then to send a proportionate inhibition to the Purkinje cell. The function computed by these cells is not modelled explicitly. It is assumed that to each Purkinje cell they furnish inhibition of magnitude

$$I = (f_3 \times P)/K_{BS}, \tag{6}$$

where P is the total excitatory input from the sampled parallel fibres (and so varies with different input patterns), K_{BS} is a constant which is equal to the ratio of the number of parallel fibres sampled by a basket or stellate cell to the number of parallel fibres sampled by a Purkinje cell, multiplied by the number of basket and stellate cells inhibiting a Purkinje cell. f_3 is a constant of value slightly less than 1.0, its value being obtained empirically, as explained below. In this way, each Purkinje cell will receive an inhibition (summed over many basket and stellate cells) which is just less than the total excitation in the parallel fibres. This will result in the Purkinje cell only firing when the vast majority of excited parallel fibres have activated synapses. Adjusting the value of f_3 adjusts the acceptable difference from a learned pattern in order to still respond to it.

(*c*) *Rules for weight changes*

The only synapses at which learning is assumed to occur are those between the parallel fibres and the Purkinje cells (Marr 1969). These all have a weight of zero initially and then are increased to a weight of one when both the presynaptic (granule) and post synaptic (Purkinje) cells are excited (Hebb 1949).

5. SIMULATION RESULTS

The simulations that were carried out had three aims.

1. To establish the values of the parameters that were not yet specified. Principally, these are the values

of f_1 and f_2, for determining the inhibition to be applied by the Golgi cells and that of f_3, which is needed for setting the threshold on the Purkinje cells in discrimination mode.

2. To see if the implementation of Marr's theory would work at all, and if so would work in the manner which we envisaged. In particular, we wanted to test the extent to which recoding of the mossy fibre input into the granule cells, together with inhibition by the Golgi cells, would improve performance by increasing the capacity of the net while at the same time minimizing the recognition of subsets and similar patterns.

3. To determine, using simulations, the capacity of Marr's model (or at least the adaptation of it which was deemed necessary in the light of the anatomical constraints that he did not consider), with the performance he calculated analytically for the original model.

(*a*) *Preliminaries*

In all the tests described below, each mossy fibre pattern was presented to be learned more than once, to simulate the effect that differing amounts of external excitation can have on the inhibition supplied by the Golgi cells. Each mossy fibre pattern was presented to be learned nine times, with external excitations differing from the internal figures by -5.0%, -3.75%, -2.5%, -1.25%, 0.0%, $+1.25\%$, $+2.5\%$, $+3.75\%$ and $+5.0\%$.

Marr suggested that between 0.3% and 30% of the mossy fibres may be active in any one event; we used a narrower range of between 2% and 20%. Therefore the term 'random mossy fibre pattern' below refers to a pattern in which the individual fibres are randomly turned on with a probability randomly chosen between 2 and 20%, which is kept constant for all the fibres of that pattern.

The simulations were written in C and run on a Sun-4 workstation, rated at 12 MIPS and with 24 Megabytes memory. Generation of the structure took 50 h of cpu time. It took approximately 2 min of cpu time to perform the calculations for a single presentation of a mossy fibre pattern, in both storage and discrimination mode. The step that took the most computer time was the computation of the inhibition supplied by the Golgi cells.

(*b*) *Mossy fibre to granule cell mapping*

The first tests look at the mapping from the mossy fibre pattern to the inhibited granule cell pattern to see how well the Golgi cells are able to regulate that mapping.

(i) *Golgi cell sampling of parallel fibre excitation*

Figure 16 shows the accuracy of the Golgi cells in sampling the activity in the parallel fibres. The estimates are usually too high because the cells use the higher of the two estimates from their dendritic fields. Despite the effect of this error and also of the error from the sampling of external cells, it can be seen that the Golgi cells still estimate the parallel fibre excitation to within 5%.

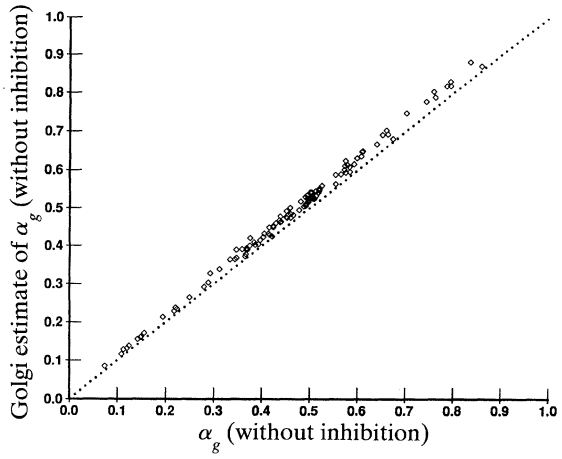

Figure 16. Golgi cell estimation of what the average activity α_g of the granule cells would have been without inhibition by the Golgi cells, versus an exact calculation of what it would have been.

(ii) *Mossy fibre to granule cell mapping*

The values of the constants f_1 and f_2 in equation 5, which determine the amount of Golgi cell inhibition as a function of the estimated level of granule cell activity before inhibition, were determined empirically to be 2.25 and 0.60. Figure 17 shows the result of the Golgi inhibition. As can be seen, the average granule cell activity after inhibition is $\approx 1.0\%$ and there is a roughly monotonic relationship between granule cell activity α_g and mossy fibre activity α_m. The slight 'saw-tooth' effect is due to the fact that Marr assumes integer values in the granule cells, whereas the Golgi cell inhibition is necessarily analogue.

Note that for 13 000 mossy fibres and 200 000 parallel fibres, the conditions in § 3*d* give rise to the two

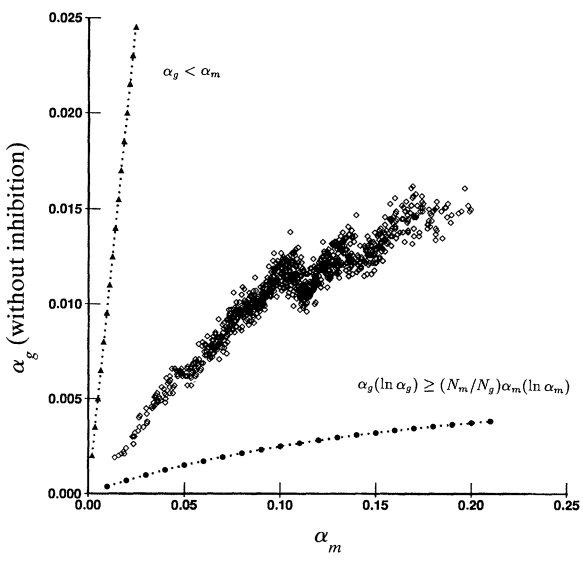

Figure 17. Golgi cell regulation of the granule cell activity level (α_g). The lower dotted line (circles) corresponds to the first condition of § 3*d*, the upper dotted line (triangles) is the second condition of § 3*d*.

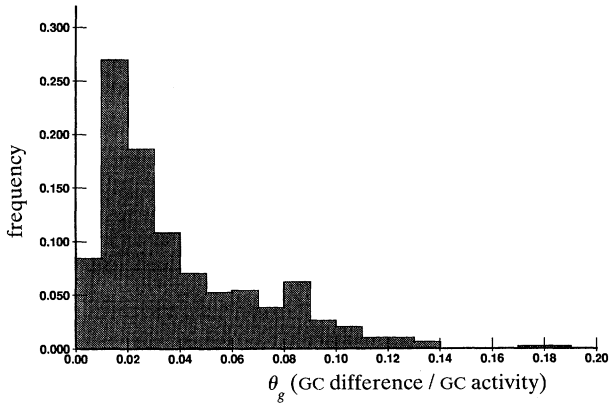

Figure 18. Variation in granule cell patterns from the same mossy fibre pattern (due to different Golgi cell inhibition). The frequency measures how often two granule cell (GC) patterns with the relevant difference measure are generated from the same mossy fibre pattern. The difference measure is the number of granule cells at which activity differs divided by the mean number of active granule cells in the two patterns (see § 3*d*).

dotted lines in figure 17. As can be seen, the relationship between α_g and α_m is properly balanced between the two criteria.

The variation in the granule cell patterns due to a varying external Golgi cell input (for the same mossy fibre pattern) is shown in figure 18. It is seen that about 50% of the granule cell patterns are less than 3% different from each other, whereas about 90% are less than 8% different. There was a variation of up to $\pm 5\%$ in the external mossy fibre and granule cell excitations.

In § 3*c* it was hypothesised that another property of the mossy fibre to granule cell complex connections would be that two similar mossy fibre patterns would produce two granule cell patterns which are less similar. Figure 19 shows that this does indeed happen.

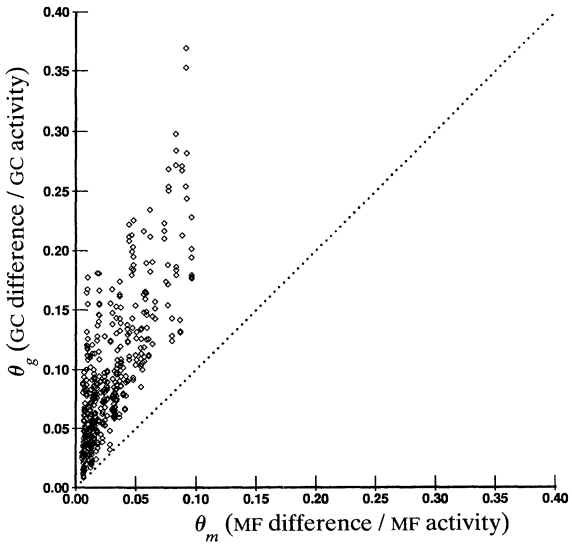

Figure 19. Expanding the difference between pairs of input patterns as they pass from the mossy fibres (MF) to the granule cells (GC). θ_g is seen to be always greater than θ_m.

(*c*) *Basket and stellate sampling of the parallel fibre excitation*

These cells were only modelled implicitly and their behaviour is highly dependent on the assumed difference between internal and external excitations (figure 20).

(*d*) *Discrimination mode*

Calculating the capacity of the net involves two steps: (i) calculating the value of f_3 (by fixing the acceptable error rate for recognition of learned contexts); and (ii) calculating the capacity of the net (by fixing the acceptable error rate for recognition of unlearned contexts).

Both error rates are set to be 1% in accordance with the figure adopted by Marr. Thus 1% of learned contexts are not responded to, and 1% of unlearned contexts are responded to.

In the first step, the value of f_3 is found empirically by storing in the net 540 patterns (sets of the nine variants of 60 basic patterns) and then finding the highest value of f_3 for a 1% error rate. This yielded a value of $f_3 = 0.935$.

(i) *Capacity*

Once the value of f_3 was set, the capacity of the net is then determined in the second step by finding how many contexts can be stored before the acceptable error rate of 1% is exceeded for unlearned patterns. As figure 21 shows, the full net was found to have a capacity of between 60 and 70 contexts. After having learned 60 associations, 22.8% of the parallel fibre to Purkinje cell synapses had been modified. The purpose in stating the capacity we found is not so much to give a realistic estimate of the capacity of the cerebellum as to show that each Purkinje cell can plausibly be expected to store many associations (i.e. that it is feasible that the Purkinje cell is the equivalent of an output line in an associative net).

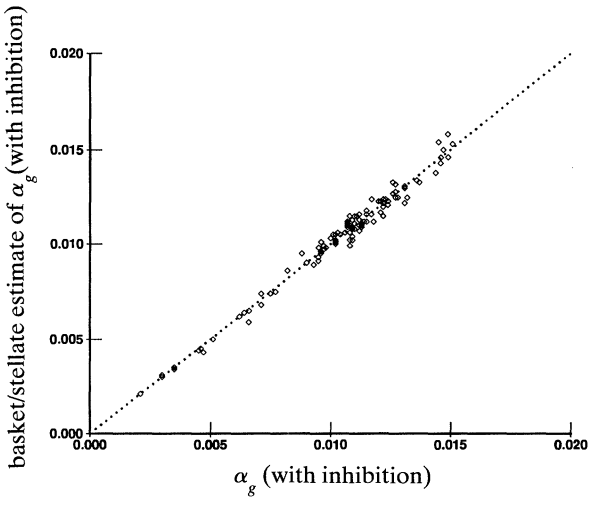

Figure 20. Basket and stellate cell estimation of the average excitation of the granule cells (α_g) versus the true value of α_g.

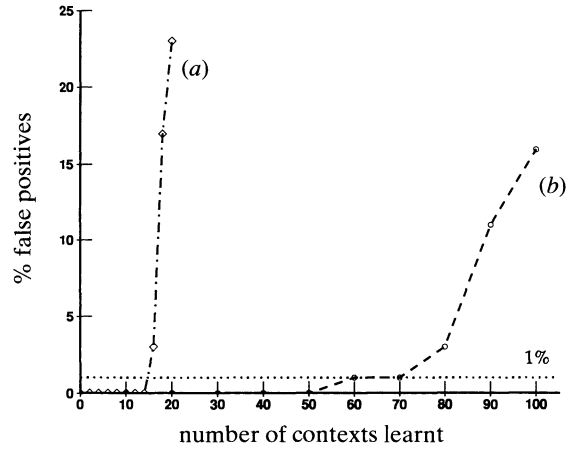

Figure 21. (*a*) Capacity of a net without granule cells and Golgi cells; (*b*) capacity of the whole net. Both capacities are determined by fixing the error rates due to both incorrect recognition of unlearned contexts, and incorrect rejection of learned contexts, at 1%. It can be seen that the whole net can learn many more associations then the simplified net.

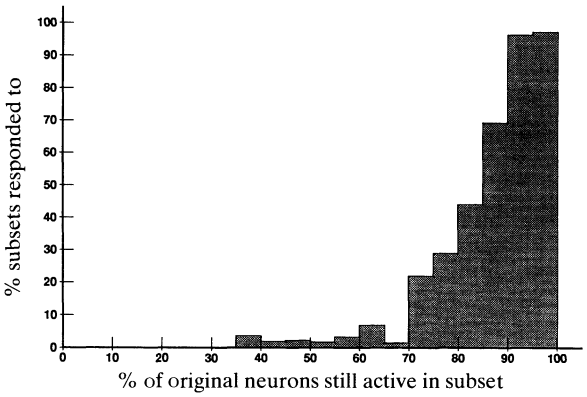

Figure 22. Recognition of subsets of learned patterns. Percentage of subsets erroneously recognized versus the percentage of the original learned pattern remaining in the subset.

(ii) *Responding to subsets and discriminating between similar patterns*

In § 3*c* it was hypothesized that one property of the mossy fibre to granule cell complex transformation would be to 'scramble' input patterns so that subsets of learned mossy fibre patterns were not automatically responded to. Figure 22 shows that subsets of learned patterns in which fewer than 70% of the neurons originally active were still active hardly ever elicited a response.

Figure 23 shows how many false positives (incorrect responses to unlearned patterns) occur when the input patterns are similar to already learned ones. The figure shown is the result of presentation of 1000 patterns.

In both tests, the results depend on how many synapses have been modified (how many associations have been learned) and also on the value of f_3 in § 4*b*, which specifies the tolerance for patterns that are similar to learned ones. In both cases, 60 associations had been learned and a value of 0.935 was used for f_3.

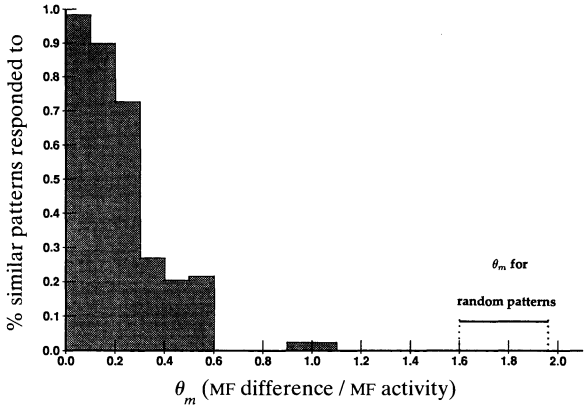

Figure 23. Recognition of patterns similar to ones already learned. Percentage of similar patterns erroneously recognized versus a measure of their difference from the learned pattern. The measure of difference (θ_m) is the number of mossy fibres at which activity differs divided by the mean number of active mossy fibres in the two patterns. The bar shows the expected range of values of θ_m for pairs of randomly chosen patterns with activity ranging from 0.02 to 0.2.

(iii) *Comparisons with a simpler structure*

The most important advantage postulated for the mossy fibre to granule cell set of synaptic connections, together with the Golgi cell inhibition, is that of increasing the capacity of the net. To test this we compared the capacity of the full net of figure 4 with that of a comparable net but without granule and Golgi cells, so that mossy fibres synapse directly onto the Purkinje cells and the basket and stellate cells sample the mossy fibres. The second net is the biological equivalent of the standard associative net of figure 8.

As already described, the highest value of f_3 for the full net, that obtained a 1% error of omission, was 0.935, resulting in a capacity of 60–70 contexts.

For the simplified system, a 1% error rate for response to learned input patterns led to a value for f_3 of 0.92, and an identical error rate for response to unlearned input patterns led to a capacity of ≈ 15 contexts (see figure 21).

This demonstrates that the complex expansion from mossy fibres to granule cells, together with the subsequent inhibition by the Golgi cells, does increase the number of associations that can be stored and retrieved reliably by the network. This is mainly because the expansion to a larger set of granule cells allows more sparsely coded patterns.

6. DISCUSSION

In the introduction we stated that this work consisted of three main parts: a concise but clear explanation of Marr's theory of the cerebellum, a description of how to simulate the neural structure of the cerebellum, and some insights into Marr's theory gained by forcing it to be implemented in our simulation. In this discussion we summarize these three parts and then proceed to make a claim about the importance of Marr's ideas.

(*a*) *Explanation of Marr's theory*

(i) *Functions of the different cells*

In § 3 we explained Marr's 'theory of cerebellar cortex' by means of an extended analogy with the associative net. The different cells in the cerebellum were compared with the following parts of the extended associative net: (i) mossy fibres: first layer input lines; (ii) granule cells: second layer input lines; (iii) Purkinje cells: output lines; (iv) climbing fibres: 'teacher' lines;(v) basket and stellate cells: threshold setters; and (vi) Golgi cells: regulators of second layer input activity.

It was also hypothesized in § 3, and shown as reasonable in § 5, that the existence of the Golgi and granule cells and the complex mossy fibre to granule cell connections can be explained by the need to improve the performance of the basic associative net in three ways: capacity, rejection of subsets, and rejection of similar patterns.

(ii) *General problems with Marr's theory*

Up to this point the only drawbacks to Marr's theory that we have discussed have arisen out of the problems we encountered in trying to map Marr's fairly mathematical model onto our simulated neurobiology. It is worth mentioning here that there are other, more general, problems with his theory.

1. Associations cannot be unlearned. Because the synapses in the model can only be turned on, associations between input and output can never be overriden. A situation–action link, once formed, must then remain in place for ever.

2. Binary synapses do not allow for a very refined sort of learning. Marr's theory presupposes that the type of learning that we want from the cerebellum is a sort of 'photographic recognition', whereby an input pattern is only recognized if nearly all its constituent elements (analogous to 'pixels') are exactly the same as in the learned 'photo'. There is no possibility of any learning in which certain portions of the input can be effectively ignored as not relevant, whereas others can be given greater importance if they are more significant in deciding whether the context should be recognized.

To illustrate this concept, consider a Purkinje cell which is responsible for the control of a muscle that causes the arm to flex at the elbow, and for which two parts of the input consist respectively of information about the degree of rotation of the elbow and information about the degree of flexion of the ankle (if two such inputs would indeed converge on the same Purkinje cell). In this case, the latter part of the mossy fibre context would not be very relevant to any decisions about whether the Purkinje cell should fire, whereas the former part of the mossy fibre context would be of considerable importance. Marr's learning rules, involving only binary weights on the parallel fibre to Purkinje cell synapses, are not able to produce varying emphasis on the different parts of the input.

3. In general, Marr's use of binary synapses and integer excitations is biologically dubious. Albus (1981) comments:

> Neurons are *not* binary devices, and the brain is not a digital computer. The all or nothing character of the action potential does not mean that the neural signal is a Boolean variable. The action potential is simply an encoding mechanism that the brain uses for transmitting analog variables over long distances.

4. Experimental evidence seems to suggest that synaptic values start off high in the cerebellum and are then decreased with the conjunction of parallel fibre and Purkinje cell activity (Gilbert & Thach 1977), rather than being increased from low to high, as assumed by Marr. The Purkinje cell may learn when to pause its inhibition (to 'disinhibit') rather than when to fire.

Although the problems just outlined should make us feel sceptical about some of the details of Marr's theory, they should not make us feel too dubious about the more fundamental aspects of Marr's theory (that the cerebellar cortex associates patterns in its input with patterns in its output, and that the component cells function as outlined in § 6*a*(i)). The basic associative net can be modified to work with analogue excitations and synapses, and different synaptic rules which involve depression as well as potentiation of synapses can be used (Albus 1971).

(*b*) *Simulating the cerebellum*

We have demonstrated a simulated full-size model of what is generally thought of as the building block of cerebellar circuitry: the cells and synaptic connections associated with a single Purkinje cell. In our model most of the cells and processes could be represented explicitly and in a form that captures their spatial arrangement. In particular, the model reproduces the probabilistic aspects of cerebellar structure. Cells are positioned stochastically and have varying numbers of dendritic and axonal connections which are positioned randomly around the cell body.

The simulation consisted of a population of 13 000 mossy fibres that innervate 200 000 parallel fibres under the regulation of 100 or so Golgi cells. The parallel fibres then synapse with a single Purkinje cell. The parallel fibres also pass excitation to the Purkinje cell by way of 40 basket and stellate cells. The Purkinje cell also receives input from one climbing fibre.

We have used this simulated structure to show how it can embody the Marr theory of the cerebellum as an associative learning device. More generally, this model can be regarded as a simple building block for associative memory. But when it is applied to other structures the way it will work in detail will depend heavily on the specific numerical relationships of the structure under consideration.

It will be possible to use this structure to test out the performance of other proposed models of the cerebellum. A case in point is that based on the theory due to Albus (1971), whereby Purkinje cells are to be taught by error correction. In this theory, synapses are analogue rather than digital, and the proposed learning mechanism involves depression of parallel fibre

synaptic weights rather than strengthening, as well as modification of the parallel fibre to basket and stellate cell synapses.

Even in a model of this precision, the values of many of its parameters had to be guessed. Given the current state of techniques, it is possible that many missing pieces of neuroanatomical data can now be obtained. Particularly helpful would be more exact estimates of the following parameters: (i) the size and shape of Golgi cell processes; (ii) the numbers of mossy fibres synapsing onto each Golgi cell's descending dendrites; (iii) the numbers of parallel fibres synapsing onto each Golgi cell's ascending dendrites; (iv) the numbers of granule cell dendrites contacted by each Golgi cell's axons; (v) the number of parallel fibres synapsing onto the dendrites of each basket or stellate cell; and (vi) the average number of mossy fibre axon terminals in each cluster.

(*c*) *Insights into Marr's theory*

When we implemented Marr's theory in our simulation of the cerebellum we encountered some places where Marr's neuroanatomical assumptions clashed with what is now known. A general comment is that although Marr's implementation was spelled out in much detail, in some respects it was found to be inadequate.

1. A representation of the mossy fibre – granule cell – Purkinje cell pathway was made that is self-consistent, relies on plausible assumptions for cases where data is unavailable and can be seen to be a natural implementation of Marr's basic idea. One place in which our interpretation differs from that of Marr is that we found that physical constraints dictate that the 6000 mossy fibres known to influence one Purkinje cell directly be increased by edge effects to 13 000 rather than to 7000, as suggested by Marr.

2. Marr's interpretation of the anatomy of the Golgi cells contravenes many pieces of anatomical data and seems to be biologically untenable. In our model, Golgi cells sample mossy fibres that are outside the 13 000 leading to the given Purkinje cell. This means that the same mossy fibre pattern will, at different times, cause activity in slightly different populations of granule cells. It has the implication that there is a stochastic element in the mossy fibre to granule cell transformation, in contrast to Marr's assumption that a given mossy fibre input to a Purkinje cell should be carried there by parallel fibre activity which is determined by that input alone.

(*d*) *The importance of Marr's theory*

As has just been discussed in §§ 6*a*(ii) and 6*c*, there are problems with some of the details of Marr's theory. However, none of these problems is fatal to the fundamentals of the theory. We should not look to Marr for an exact prescription as to how the cerebellar cortex works, but yet we should note the plausibility and aptness of the analogy with the associative net and the functions he ascribes to the different cell types. No other theory links the structure and function of the cerebellum in such a convincing and explicit fashion.

Some years after having written his 'Theory of cerebellar cortex', Marr himself said 'I shall be very surprised if my 1969 [cerebellum] or 1971 papers turn out to be very wrong.' (Vaina 1991). Thach recently commented 'Certainly some elements of Marr's work may require modification. Yet, a growing number of network theoreticians and experimental neuroscientists appear to like the ideas, and to anticipate their being proven to be essentially and substantially correct.' (Vaina 1991). We agree.

We thank Jay Buckingham, Marcus Frean, Peter Dayan, Kate Jeffery, Carl Rasmussen, Neil Thacker and Patrick Courtney for commenting on drafts of this paper, and John Hallam for assistance with the work itself. This project was supported by the SERC (grant no. 88411480) and the MRC (Programme Grant no. PG8514914).

REFERENCES

Albus, J.S. 1971 A theory of cerebellar function. *Math. Biosci.* **10**, 25–61.

Albus, J.S. 1981 *Brains, behaviours and robotics.* BYTE Publications. Lawrence Erlbaum Associates.

Baron, R.J. 1987 *The cerebral computer.*

Braitenberg, V. 1961 Functional interpretation of cerebellar histology. *Nature, Lond.* **190**(4775), 539–540.

Carlson, N.R. 1977 *Physiology of behaviour.* Boston: Allyn and Bacon.

Changeux, J.-P. 1985 *Neuronal man – the biology of mind.* Oxford University Press.

Churchland, P.S. 1986 *Neurophilosophy: toward a unified science of the mind/brain.* Massachusetts Institute of Technology Press.

Eccles, J., Ito, M. & Szentagothai, J. 1967 *The cerebellum as a neuronal machine.* Berlin: Springer–Verlag.

Fox, C.A., Hillman, D.E., Siegesmund, K.A. & Dutta, C.R. 1967 The primate cerebellar cortex: A Golgi and electron microscope study. *Prog. Brain Res.* **25**, 174–225.

Fujita, M. 1982 Adaptive filter model of the cerebellum. *Biol. Cyber.* **45**, 195–206.

Gilbert, P.F.C. 1974 A theory of memory that explains the structure and function of the cerebellum. *Brain Res.* **70**(1), 1–18.

Gilbert, P.F.C. & Thach, W.T. 1977 Purkinje cell activity during motor learning. *Brain Res.* **128**, 309–328.

Gilman, S., Bloedel, J.R. & Lechtenberg, R. 1981 *Disorders of the cerebellum.* F. A. Davis Company.

Hebb, D. 1949 *The organization of behavior.* New York: Wiley.

Ito, M. 1982 Mechanisms of motor learning. In *Competition and cooperation in neural nets.* Berlin: Springer–Verlag.

Ito, M. 1984 *The cerebellum and neural control.* New York: Raven Press.

Kanerva, P. 1984 *Self-propagating search: a unified theory of memory.* Centre for the Study of Language and Information, Stanford University.

Kohonen, T. 1978 *Associative memory.* Berlin: Springer-Verlag.

Lippmann, R.P. 1987 An introduction to computing with neural nets. *IEEE ASSP Mag.* **3**(4), 4–22.

Llinás, R.R. 1973 The cortex of the cerebellum. *Scient. Am.* **232**(1), 56–71.

Marr, D. 1969 A theory of cerebellar cortex. *J. Physiol.* **202**, 437–470.

Marr, D. & Blomfield, S. 1969 How the cerebellum may be used. *Nature, Lond.* **227**, 1224–1228.

Palay, S.L. & Chan-Palay, V. (eds) 1982 *The cerebellum: new vistas*. Berlin: Springer–Verlag.

Pellionisz, A. & Llinás, R. 1982 Tensor theory of brain function: the cerebellum as a space-time metric. In *Competition and cooperation in neural nets* (ed. S. Amari & M. A. Arbib), pp. 394–417. Berlin: Springer–Verlag.

Pellionisz, A.J. 1986 David Marr: A theory of the cerebellar cortex. In *Brain theory* (ed. G. Palm & A. Aerten), pp. 253–257, Berlin: Springer-Verlag.

Steinbuch, K. 1961 Die Lernmatrix. *Kybernetik*, **1**, 36–45.

Szentagothai, J. & Arbib, M.A. 1975 *Conceptual models of neural organization*. Massachusetts Institute of Technology Press.

Thompson, R.F. 1990 Neural mechanisms of classical conditioning in mammals. *Phil. Trans. R. Soc. Lond.* B **329**, 161–170.

Vaina, L.M. (ed.) 1991 *From the retina to the neocortex: selected papers of David Marr*. Birkhauser.

Willshaw, D.J. 1971 *Models of distributed associative memory*. Ph.D. thesis, University of Edinburgh.

Willshaw, D.J., Buneman, O.P. & Longuet-Higgins, H.C. 1969 Non-holographic associative memory. *Nature, Lond.* **222**(5197), 960–962.

Received 7 November 1991; accepted 21 January 1992