BIOLOGICAL CONSTRAINTS ON CONNECTIONIST MODELLING

Simon J. Thorpe and Michel Imbert,
Institut des Neurosciences,
Departement des Neurosciences de la Vision,
Université Pierre et Marie Curie,
9, Quai St. Bernard, 75005, Paris, FRANCE.

ABSTRACT

Many researchers interested in connectionist models accept that such models are "neurally inspired" but do not worry too much about whether their models are biologically realistic. While such a position may be perfectly justifiable, the present paper attempts to illustrate how biological information can be used to constrain connectionist models. Two particular areas are discussed. The first section deals with visual information processing in the primate and human visual system. It is argued that speed with which visual information is processed imposes major constraints on the architecture and operation of the visual system. In particular, it seems that a great deal of processing must depend on a single bottum-up pass. The second section deals with biological aspects of learning algorithms. It is argued that although there is good evidence for certain coactivation related synaptic modification schemes, other learning mechanisms, including back-propagation, are not currently supported by experimental data.

Keywords : Vision, Learning Algorithms, Back-propagation, Hebb, Parallel Processing, Multi-layer architectures, Feedback, Coding, Context, Temporal constraints.

1. INTRODUCTION

Connectionist models can be used in two rather different ways. The aim of the first approach is the development of artificial intelligent systems. If this is the aim then biological realism is not really an important design factor - one can take the view that although "neurally-inspired", the actual architectures and learning algorithms used can be fundamentally different from those used in real brains - the only real criterion for progress is "Does is work?" Alternatively, connectionist

models can be specifically used to further our understanding of how real nervous systems function. Obviously, in such a case, a high priority should be placed on making the model as biologically reasonable as possible. This sort of direct modelisation of neural systems was previously rather rare, but a number of examples have appeared in the literature in the last year or so (see for example, Lehky and Sejnowski (1988), Zipser and Anderson (1988), Sejnowski, Koch and Churchland (1988), Thorpe and Pouget (this volume)).

In principle, there is no reason why these two different ways of using connectionist modelling could not be used relatively independently. Nevertheless, it would be unfortunate if applications-oriented research workers were ignorant of the progress being made in understanding the function of real brains. After all, the human brain is the product of millions of years of trial and error - the solutions adopted as a result of intense natural selection may not be the only way of solving problems, but at least they work! On the other hand, researchers working on biological systems can benefit from the insights obtained by those involved in artificial intelligence research.

Biological data can potentially provide useful information about a number of different computational issues. One important area is coding. Connectionist coding schemes can be distributed, local or something in between (Hinton and Anderson, 1981; Smolensky, 1988; Rumelhart and McClelland, 1986; Feldman, 1985), and biological data may well help in defining those situations in which it may be advantageous to use one or other. Another important area concerns the precision with which such values as synaptic weight strengths should be coded. Current hardware implementations of neural networks use 1, 2, 4, 8 or even 16 bit precision. Biological data could help those involved in designing such systems to decide what level of precision to use. However, in this chapter, we would like to concentrate on two other areas in which biological constraints may be important. The first area concerns the way in which information is processed by the visual system. There are currently a large number of suggestions as to how visual information is handled, involving various degrees of "bottom-up" and "top-down" processing, mechanisms akin to relaxation and so on. It will be argued that the speed with which visual information is processed, together with information about the anatomical organization of the visual system and the activity of neurons in the visual system, provide heavy constraints on the types of models which are biologically feasible.

The second topic concerns the sorts of learning algorithms used by nervous systems. Various possibilities have been proposed, including variations on Hebbian plasticity, the Perceptron learning algorithm, back-propagation and the Boltzmann algorithm. Neurophysiological, psychological and theoretical reasons will be

presented for thinking that not all such mechanisms are likely to be implemented in biological "wet-ware".

Before discussing these two problems in detail, it may be worth first characterizing in a bit more detail the sorts of selection pressures that were at work during the evolution of the human brain. Natural selection has probably favoured those variations which improve the animsls ability to generate appropriate behavioural responses to environmental situations. However, not all such solutions are likely to be selected. One particularly important factor is speed. If we consider two animals - both of which have visual systems capable of detecting the presence of a predator - but one of which can reach a decision a few milliseconds faster than the other, even such a small difference will be rapidly amplified in the survival rates of their respective offspring. This speed constraint will have been particularly important, given the relatively slow nature of information transmission by individual neurons.

A second factor which will have been of vital importance is size. Again, let us consider two animals, both of which can detect the presence of a predator. Let us assume that both have visual systems that can come to a decision after the same amount of processing time. However, suppose that one can solve the problem with a visual system containing 1,000,000 neurons, whereas the other needs 10% more. It has to be remembered that 10% more neurons means that the animal has to provide 10% more energy for the metabolic requirements of this extra tissue, and there will also be a cost related to the extra weight that must be carried. Thus all other things being equal, the 1 million neuron solution will have a significant advantage. Calculations have shown that even minute differences in reproductive success can result in significant selection over a large number of generations (Dawkins, 1987).

Clearly, the constraints that apply when designing an artificial vision system are not usually so tight. An increase of 1% in the time taken to reach a decision would not be considered significant for most engineering applications, and if 10% more components were used, this would increase the price slightly, but at least in the absence of competition from Taiwanese clones, such extravagance can usually be tolerated!

2. PROCESSING

Jerry Feldman has pointed out a major biological constraint on information processing by animals and humans (Feldman and Ballard, 1982; Feldman, 1985a, Feldman, 1985b. Sophisticated behavioural responses can often be generated only a few hundred msec after a change in the environment. Since neurons can only generate spikes at about one every 5 msec, even when firing maximally, Feldman concluded that this puts a limit of roughly 100 computational steps between input and output. In this section, we would like to see what happens if we limit the discussion to the sensory processing component.

2.1 Information processing by the brain

The human brain has specialized structures for processing information from various sensory modalities - vision, audition, touch, smell and taste. For each of these modalities, we can present a sensory stimulus and ask how long it takes for it to be processed. In general, processing will be taken as meaning the time taken for the stimulus to be identified. In a number of respects, the question is easier to pose in the case of vision. Presenting olfactory and gustatory stimuli at a precise moment is difficult, and so measuring latencies is problematic. In the case of audition, much auditory information is by its very nature temporally structured and so presenting an auditory stimulus such as a word obviously takes a certain time - so here again measuring processing time becomes difficult. Touch is also often used in a dynamic way - we often explore an object actively in order to identify it by touch. However, in the case of vision, it is relatively easy to present even complex visual patterns at a precise point in time and see how long processing takes.

This is the principle reason why we will restrict our discussion of sensory processing to the visual modality. However, it is also the case that vision has for various reasons been more intensively studied in the past. The primary question that we will address concerns the time taken by the visual system to process an image. Both Human Experimental Psychology and Neurophysiology provide data relevant to this issue.

2.2 Psychological data on visual processing time.

Visual identification is so easy for humans that we often fail to realize just what a difficult task it is. Imagine flicking through a magazine full of photos. Even if the photos are of widely varying objects and scenes, we usually have no trouble in identifying what they are. How long does this identification process take? A variety of different approaches have been used in Psychology, and these will be briefly reviewed

in this section (see also Riddoch and Humphreys, 1988; Humphreys and Quinlan, 1988). Other neurophysiological evidence will be discussed in the next section.

Naming time. Perhaps the most obvious source of information is simply to measure the time it takes people to name a picture of an object or scene. This approach goes back more than 100 years to the studies of James McKeen Cattell (1886a,b), whose findings have been expanded on by more recent studies (e.g. Oldfield, 1966; Potter and Faulconer, 1975; Intraub, 1979; Carr et al, 1982; Henderson et al, 1987). Typically, it takes about 750 msec to name an object. However, the time taken depends on the object in question - an apple, for example, can be named in 600 msec (Intraub, 1979), whereas less familiar objects, such as a xylophone can take as long as 1200 msec (Oldfield, 1966). The problem with such an approach is that we measure not only the time taken by the visual processing, but also the time needed to find the right word to describe the object and the time taken to generate the verbal response. We still need to know what proportion of the naming time is taken up by the visual processing per se.

Minimum presentation time. Another possibility is to measure the minimum presentation time required for an object to be identified correctly. Such studies have shown that in most cases, a presentation lasting 10 to 15 msec is sufficient (Oldfield, 1965; Intraub, 1979). However, this is only the case if the picture is followed by a blank field. If the presentation is followed by the presentation of a complex visual mask, the time required increases to 100 to 120 msec, but even in this case, it is difficult to be sure that visual processing does not continue after this time.

Rapid serial visual presentation. In 1975, Molly Potter, working at M.I.T., introduced an new technique for studying visual processing (Potter, 1975; Potter, 1976). She presented sequences of 16 colour images, with a presentation time varying between 113 and 333 msec per image. Subjects were instructed to press a button as soon as they saw a particular image (e.g. "A boat on a beach"), which could be situated anywhere in the sequence. The results showed that even with the shortest presentation times, the subjects were able to find more than 60% of the targets, leading Potter to conclude that even with presentations of only 113 msec, a great deal of visual processing was possible.

More recently, Helene Intraub has used variations on Potter's technique (Intraub, 1980, 1981). Instead of verbally specifying the target ("a baby reaching for a rattle"), the target was either specified by giving a category ("press when you see an animal") or a negative category ("press when you see something that you don't find in a house"). With presentation times of 114 msec per image, performance was somewhat lower than in Potter's original experiments, but the subjects still managed to find 46% of the targets when the category was specified and 35% when a negative category was used.

There is, however, a complication in the interpretation of such experiments. The fact the subjects are looking for a particular visual target means that in some way the visual system could be primed to look for particular visual features. What would happen if the subject had no contextual information about the type of image that was going to be presented?
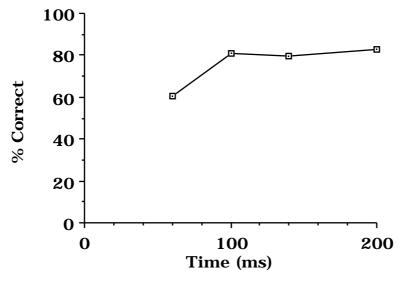

Figure 1 : The influence of exposure time on the ability of subjects to identify complex images. A pair of images was presented, and the subject required to identify both. The probability that the subject could correctly identify the first image is plotted as a function of presentation time for those trials on which the second image was also correctly identified. (Thorpe, 1988, and in preparation).

Identification of rapidly presented images. To try and answer this question, we have recently developed an experimental paradigm designed to test the performance of the visual system under conditions in which the effect of context is reduced to a minimum (Thorpe, 1988). We have digitized a large stock of pictures (faces of famous people, paintings, everyday objects, advertising logos, cartoon characters etc), chosen to be a varied as possible, but at the same time being well-known to the majority of people. Using an image processing system, we can present a pair of such images, one immediately after the other to naive subjects. Their task is to name both the images ("Charlie Chaplin", "Concorde", "The Eiffel Tower", etc.). The duration of the images was varied systematically (60, 100, 140 or 200 msec), and the effect on the percentage of images that could be successfully identified determined. Identification of both the first and second images was remarkably robust, averaging 80%, but perhaps the most interesting observation came from looking at the ability of subjects to identify the

first image, but taking into account only those trials on which they successfully identified the second. The results of this analysis are illustrated in Figure 1.

It can be seen that reducing the presentation time from 200 to 100 msec had no significant effect on identification, which remained constant at around 80%. Only when presentation time was reduced to 60 msec was there any real effect on performance, and even then, the effect was by no means catastrophic - 60% correct is still a remarkably good level of performance, given the complexity of the visual processing needed to identify the images in question.

Thus, even when the subjects had no advance information about the type of image to be presented, and when they had to process another similar image immediately afterwards, 100 msec presentation times are sufficient for accurate identification of real everyday images.

2.3 The number of identifiable objects and scenes

This level of performance is particularly impressive, especially when one bears in mind the number of such objects and scenes that can be identified in this way by the average human subject. Although some images and scenes need considerable time to be interpreted, a great many objects and scenes can be identified very rapidly and without effort. Is it possible to say how many? Clearly, this figure should include familiar objects ("can-opener", "telephone", "hair-dryer" etc), animals and plants ("tiger", "carnation", "spider" etc.), geographical locations ("The empire state building", "Big Ben" etc.), signs ("MacDonalds", "Mercedes" etc), as well as faces ("Ronald Reagan", "Michael Jackson", etc.). Biederman (1987) estimated that we can name approximately 30,000 different objects. This estimate was based on a sample of the entries in Webster's dictionary, in which he counted the number of words that corresponded to objects which could be easily identified visually. Thus he counted penguin and ostrich, since they were clearly distinguishable, but not garden warbler, which, at least for him, was not sufficiently different from other types of bird.

Nevertheless, the value of 30,000 visually identifiable objects proposed by Biederman is certainly an underestimate, for two reasons in particular. Firstly, if instead of counting entries in a dictionary, you use a large encyclopedia, you will count not only simple words such as "chair", "hammer" and "ostrich", but also specific terms such as "Eiffel tower", "Mona Lisa" and "Adolf Hitler", which are also easy for most people to identify. Using the Encyclopedia Britannica, we have made such an estimation and found a value of 50 to 60,000 visually identifiable entries.

The second point is that even this value does not include all the objects and images that are part of the personal repertoire of every one of us, such as the faces of our family and friends, the school we attended when we were young, the sofa in our living room and so on. It is even harder to get a clear idea of the size of this number but

it seems likely that the total number of objects and scenes that can be identified rapidly is well in excess of 100,000.

So psychological data leads us to the conclusion that with only 100 msec of presentation time, the human visual system can probably access any one of upwards of 100,000 different internal representations. But there are other neurophysiological arguments in support of the idea that visual information processing is a very rapid process.

2.4 Neurophysiological data on processing time.

In the last two decades, a number of researchers have started to analyze the visual response properties of individual neurons in different parts of the brain of awake behaving monkeys. Such studies have revealed that some neurons with quite complicated response properties can respond with remarkably short response latencies following the presentation of a visual stimulus.

Visual responses selective for food. In 1976, Edmund Rolls and his collaborators reported the existence of a population of neurons in a region known as the lateral hypothalamus which became active when the animal was looking at food (e.g. a piece of banana, a peanut, a syringe filled with glucose solution etc) (Rolls, Burton and Mora, 1976). Similar neurons have also been reported in the sheep (Kendrick and Baldwin, 1986). One of the remarkable features of these neurons is that they can respond to any visual stimulus that the animal associated with food, including initially neutral stimuli such as a white square, if such a stimulus came to signal that food was available. The latency at which such neurons became active was found to be around 150 msec (Rolls, Sanghera and Roper-Hall, 1979), and since in order to show such properties, the visual analysis of the stimulus must have been complete, the visual processing time must be less than 150 msec.

More recently, Thorpe, Rolls and Maddison (1983) described neurones in another brain region, namely the orbitofrontal cortex, which also had selective visual responses to food, but with even shorter latencies (around 120 msec). Interestingly, other neurons in the same area also showed selective gustatory responses, and a few neurons responded to both the sight and taste of particular foods. But the main point of relevance here is that highly processed visual information gets to the frontal lobe only 120 msec after the presentation of the stimulus.

Visual responses selective for faces. Other studies by Edmund Rolls, Dave Perrett and their colleagues have described a population of neurons in a part of the temporal lobe known as the superior temporal sulcus which have visual responses that are selective for faces and face features (Perrett, Rolls and Caan, 1982; Rolls, 1986; Perrett, Mistlin and Chitty, 1987). Similar neurons have also been reported in the temporal lobe of the sheep (Kendrick and Baldwin, 1987). The existence of such neurons implies

a high level of visual processing, especially when one bears in mind the observation that some of these neurons can respond of a large range of sizes, colours and so on. Nevertheless, the latency with which such cells respond is typically in the range 100 to 140 msec.

Visual responses to familiar objects. Finally, the same group reported a population of neurons in the anterior thalamus which showed highly characteristic responses during the performance of a visual recognition memory task (Rolls, Perrett, Caan and Wilson, 1982). These neurons had the remarkable property of responding selectively to familiar visual stimuli. If a novel visual stimulus (for instance a Coca-Cola bottle) was presented to the monkey, there was no response on the first presentation. However, if the same object was presented a second time, there was a large response from the neuron, even if the first and second presentations were separated in time by as much as 1 to 2 hours. Furthermore, these visual responses could be obtained with any visual stimulus - thus the same neuron could respond to faces, bunches of keys, pens and so on - the only essential characteristic was that the object should have been seen by the monkey shortly before.

The shortness of the latency of some of these neurons was particularly remarkable - 130 msec in one case. Given that the neurons were able to respond to any visual stimulus, as long as it was familiar, this would imply that most visual processing can be achieved in this time.

Selective evoked potentials in humans. Aled Jeffreys, working at Keele University in the U.K., has recently reported the existence of selective visual evoked responses in humans to faces, and other complex meaningful stimuli (Jeffreys, 1988). These potentials, which can be recorded from electrodes attached to the subject's scalp, apparently only occur to stimuli that the subject himself considers as "face-like". A variety of controls indicate that the potentials cannot be accounted for in terms of simple image characteristics such as contour, contrast and so on. The important point as far as the present argument is concerned is that such potentials have peaks at 150 to 180 msec, thus confirming the results of single unit studies in awake behaving animals.

2.5 Interim conclusion - Visual processing time

We have reviewed a variety of experimental results which relate to the speed with which visual processing is achieved. Data from psychological experiments indicate that the human visual system can handle images presented at a rate of about one every 100 msec or so. Even with such short presentation times, it would appear that human subjects can identify well in excess of 100,000 different objects and scenes. Although such experiments do not rule out the possibility that visual processing can continue even when the image is no longer present on the retina,

neurophysiological data obtained by recording the activity of single cells in awake behaving monkeys, and by recording visual evoked potentials in humans show that nevertheless, in 100 to 150 msec a great deal of visual information processing has been achieved, and that for many images, processing may even be complete.

What does such information tell us about the way in which information is processed by the visual system? Before coming to any firm conclusions, we will first briefly review some of the salient features of the visual system (Imbert, 1988).

2.6 Visual system anatomy

The pattern of light falling on the retina is sampled by approximately 100 million photoreceptors - the rods and cones. The pattern of activity of these photoreceptors is processed in the retina via a network of bipolar, amacrine and horizontal cells before being transmitted to the brain by approximately 1 million ganglion cells whose axons form the optic nerve. Clearly, the retina must be performing some sort of sophisticated data compression algorithm in order to get through the information bottle-neck of the optic nerve, since, on average, there is only one optic nerve fibre for every 100 photoreceptors, although this ratio varies dramatically across the retina - indeed the ratio is more like one-to-one in the fovea - the part used for the most fine detailed analysis of the image.

The first structure in the central nervous system to receive the visual information arriving from the retina is the lateral geniculate nucleus, which also contains approximately 1 million cells. These cells in turn project to the first cortical visual area, the striate cortex, also known as area 17 or V1. Here we see a massive expansion in the number of neurons available for processing the visual input - there are something like 500 million neurons in V1 alone.

In addition to V1, it now appears that there are a large number of other cortical areas that are also involved in visual processing, but that, at least in the primate, most visual information has first to pass via V1. These other areas include V2, V3, MT, V4 and Inferotemporal cortex. Knowledge about the anatomical organization of these different extra-striate visual areas is advancing rapidly (For reviews see DeYoe and Van Essen, 1988; Martin, 1988; Maunsell and Newsome, 1987; Van Essen and Maunsell, 1983; Zeki and Shipp, 1988), and it is now generally accepted that different cortical visual areas are to some extent specialized for particular aspects of visual processing. For example, neurons in MT (Mid-temporal cortex) are known to be particularly sensitive to stimulus movement (Hildreth and Koch, 1987) and lesions to this area and an adjacent area known as MST have been reported to result in selective deficits in motion perception and pursuit eye movements (Newsome and Wurtz, 1988). In contrast, neurons in V4 have been shown by Zeki to be particularly important in

colour processing (Zeki, 1980), an idea that would appear to be born out by the effects of lesions (Wild et al, 1986; but see Heywood and Cowey, 1988).
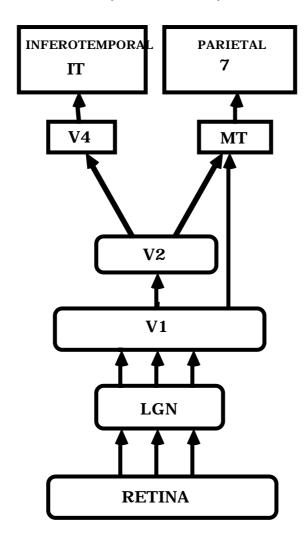

Figure 2 : Schematic diagram of the organization of the primate visual system.

At a more global level, Mishkin and his co-workers have argued that a distinction can be made between two major cortical processing streams (Mishkin, Ungerleider and Macko, 1983). The first one, involving more ventral structures and leading ultimately to the inferotemporal cortex, is involved in the identification of visual objects. The second, involving the more dorsal areas and the parietal cortex, is suggested to be more important in analyzing the spatial configuration of visual space. Clearly, from the point of view of how visual identification occurs, it is the ventral path to the temporal lobe which is the most relevant. In a previous section we mentioned the face-selective neurons described by Perrett, Rolls and others in the superior temporal sulcus (part of the temporal lobe), whose visual responses have latencies which are typically in the range 100 to 140 msec. We can ask the question

what is the shortest anatomical path by which such neurons can be activated? It would appear that in order to reach this part of inferotemporal cortex, the visual information would have to go through at least 3 intermediate cortical areas, namely V1, V2 and V4. This is because there is apparently no direct projection from V1 to V4, nor from V2 to IT (see figure 2).

A second question to ask concerns the number of synaptic stages that have to be negotiated in each cortical area. Although a definitive answer to such a question is not yet available, it does seem clear that more than one synaptic stage is involved, since the neurons that receive inputs from the preceding stage are not those which project to the next cortical area. In fact, the cortex is organized as a sandwich of horizontal layers - numbered I to VI, although subdivisions are often found. For example, in the striate cortex, layer IV has three main subdivisions, IVA, IVB, and IVC, and IVC is in turn subdivided into IVC and IVC . We know that most of the fibres arriving from the lateral geniculate nucleus synapse in IVC but that the neurons which project from striate cortex to V2 and MT are found in layers III and V. Do the neurons in IVC project directly to these output neurons? As yet, the answer is not clear - interneurons may well be necessary (Lund, 1988), but in any case, it is obvious that at least two synapses must be traversed at each cortical stage.

So, what is the minimum number of synaptic stages that have to be passed in order to get from the photoreceptors of the retina to the face-selective neurons in the temporal lobe? The answer would appear to be at least 10 - one synapse between the photoreceptors and the bipolar cells of the retina, one between the bipolar cells and the ganglion cells, one between the ganglion cells and cells in the lateral geniculate nucleus, one between geniculate cells and cells in layer IV of striate cortex, one or two between these cells and the striate cortex output cells, another 2 or 3 in V2, another 2 or 3 in V4 and at least one between the V4 output cells and the first neurons in inferotemporal cortex.

Thus, even taking the shortest known anatomical pathway involves at least 10 synaptic stages to get from photoreceptor to temporal lobe. Of course, somebody may discover a short-cut - a direct projection from the lateral geniculate nucleus to the temporal lobe for example, but for the moment, such a possibility seems completely at odds with current anatomical knowledge.
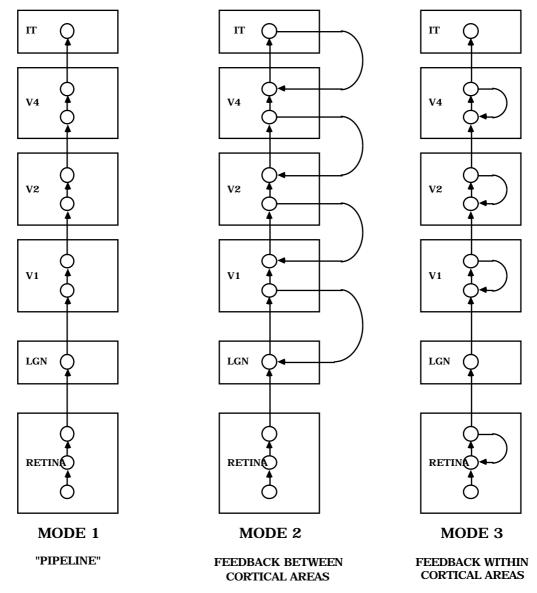
Figure 3 : Three different ways in which information may flow in the visual system.

2.7 Feed-forward and feed-back pathways.

For the moment, we have only considered the feed-forward connections which would allow information to get from the retina to the temporal lobe. In fact, recent anatomical findings stress the numerical importance of feedback pathways in the visual system. For example, there are apparently virtually the same number of fibres from V1 to the the lateral geniculate as there are running in the "normal" feedforward direction. Similar findings have been reported for the connections between cortical areas - in general, wherever there is a feed-forward projection from one area to another, one can find a substantial reciprocal backward projection. Interestingly, it appears that cells in different layers are involved in making these forward and

backward projections - and this probably is highly significant in terms of cortical information processing (Zeki and Shipp, 1988).

To what extent are these feedback pathways used in normal visual processing? In principle, one can imagine the visual system being used in a variety of different ways. Let us consider the anatomical pathway from the retina to the temporal lobe, in which at least 10 synaptic stages are involved. The simplest possibility is that the information flow is exclusively feedforward - in which case the visual system can be considered as a sort of "pipeline" processor (see Figure 3, left). Alternatively, the flow of information could involve numerous feedback loops between cortical areas (Figure 3 centre). or loops within a particular cortical area (see Figure 3, right).

Which of these different schemes actually applies during normal visual processing? To what extent can visual processing be achieved using only feedforward pathways? Are feedback loops always used? The next two sections will briefly discuss two ways in which we are currently trying to get a handle on these important questions.

2.8 Neurophysiological approaches

One way to approach the problem is to record the activity of single neurons in the visual system, present a visual pattern, and determine how the neuronal response changes as a function of time. Various types of information can be obtained using such an approach. The first thing one can learn is the latency at which neurons in different structures start to become activated after the presentation of the visual stimulus. We have already mentioned the work of Rolls and his collaborators who have shown that in the awake behaving monkey, neurons in the temporal lobe typically have response latencies of around 100 to 140 msec. Preliminary work in our own laboratory indicates that in V1, latencies are often in the range 60 to 70 msec. Such data gives some sort of feel for the sequence in which visual structures are activated.

But the temporal analysis of information processing by neurons in the visual system need not be limited to simply measuring latencies. Yves Trotter, Simona Celebrini and ourselves have recently been looking at whether the response properties of neurons in visual cortex vary in the 100 msec or so following the presentation of a visual stimulus (manuscript in preparation). We have been interested in one of the characteristic properties of neurons in the visual cortex, namely, their selectivity to stimulus orientation. Is this property already present as soon as such neurons start to respond, or is orientation selectivity something that changes as a function of time? Some models in which feedback loops are used to progressively increase orientation selectivity would require that at the moment they

start to fire following the presentation of a stimulus, the neurons should be relatively non-selective.

Although we have only recently started these investigations, the results already look promising. By analysing orientation selectivity in 10 msec time slices, starting from the moment that the neuron starts to respond, we have found that at least for some neurons, orientation selectivity is already present at the moment the neuron starts to respond, and remains virtually constant during the following 50 to 100 msec. In other words, such neurons appear to code the same information when they start responding at around 60 msec as they do 50 or 100 msec later. This would imply that at least for the stimuli used, feedback loops and processes related to relaxation do not seem to play a vital role. However, it could be that the grating stimuli used in our experiments are too simple - perhaps the situation would be different with more natural complex stimuli. Furthermore, it should be pointed out that this result is somewhat controversial - Dinse and Best (1988) and Shevelev (1987) have both reported changes in orientation preference as a function of time, although their experiments are based on recordings from anaesthetised cats, whereas our own data was obtained from awake behaving primates. On the other hand, Edmund Rolls and his collaborators (personal communication) have recently been looking for temporal changes in the selectivity functions of face-selective neurons in the temporal lobe of the awake monkey, and their results are in agreement with our own. It appears that in general, at the moment the neurons start to respond (around 100 msec) they already show the same pattern of selectivity as they will later on.

Clearly more work is required on this important question, but these preliminary results indicate that at least some neurons in the visual system are not influenced by subsequent processing and seem not to be modulated by "top-down" influences.

2.9 Temporal constraints

In addition to evidence provided by direct neurophysiological analysis, there is also a logical argument that considerable visual processing must be achieved without extensive use of feedback loops. The argument is as follows. In an earlier section we pointed out that anatomically, a minimum of 10 synaptic stages are involved in getting from the photoreceptors of the retina to the visual areas of the temporal lobe. Given that neurons in the temporal lobe have visual responses with latencies of the order of 100 msec, it is therefore inevitable that at least 10 synaptic stages must have been negotiated in this time - that is to say, an average of 10 msec per synapse. In fact, synaptic transmission can be considerably faster than this - synaptic delay is normally considered to be about 1 msec, and thus in the case of neurons which are relatively close to each other, and where axonal conduction time is not too long, it is

theoretically possible for a nervous message to get through as many as 50 synapses in 100 msec. However, V1 and the temporal lobe are not close together. In the primate, there is a distance of some 20 to 30 mm to cover, and the distance is even greater in humans. Recent estimates put the conduction speed of intracortical fibres at around 1 m/sec (Bullier et al, 1988) which presumably slows the arrival of spikes in the temporal lobe by a further 20 to 30 msec. Thus it would seem that the number of synaptic integration steps is at least 10 (the anatomical minimum) and considerably less than 50 (the neurophysiological maximum).

To realize the significance of this observation, we will have to take into account neurophysiological data on the firing rate of neurons in the visual system. Neurons normally communicate by sending a string of action potentials along their axons - information is generally thought to be transmitted by varying the interval between successive spikes. In general, firing rate varies between zero and around 100 spikes per second (i.e. one spike every 10 msec), although for a fraction of a second, some neurons can be seen to generate spikes every 5 or so msec.

The inevitable conclusion is that since the visual information must pass through a minimum of 10 synaptic stages in 100 msec, and that neurons at each stage cannot generate more than 1 or at most 2 spikes in 10 msec, **each stage must respond on the basis of 0, 1 or at most 2 spikes in the neurons of the previous stage.** It should be remembered that this calculation is based on the assumption that the visual information takes the shortest known anatomical pathway from the retina to the temporal lobe. If visual information processing involves the use of feedback loops, the temporal constraints would be even more severe, and each cell in the processing sequence would be effectively reduced to signalling binary values (i.e. 0 or 1).

Of course, we have to remember that an individual neuron does not only receive information from a single neuron in the preceding layer. There may be 1000 or more inputs converging on any particular neuron, and although the "decision" as to whether or not to generate a spike may be based on the presence or absence of a spike in each of the input axons, considerable subtlety can be obtained by combining the information from a large number of different neurons.

Nevertheless, this conclusion has a number of important implications for the way in which information is processed by the visual system. These different points will be discussed separately.

Analog coding by individual neurons. It is commonly felt that the firing rate of individual neurons can be used to code analog values such as intensity - the higher the firing rate, the higher the intensity. While this may well be the case, in the particular case of rapid analysis of information by the visual system, the fact that each stage of neuronal processing has to respond on the basis of the presence or absence of activity in the neurons of the preceding stage means that firing rate as such cannot be used to

code analog values with any real precision. However, as will be argued later, it could be that spike arrival time could be used to code such analog parameters as contrast or the probability that a particular feature is present.

Feedback Loops. If neuron A projects to neuron B which in turn can influence neuron A, and if neuron A has a maximum firing rate of one spike every 5 msec, it is clear that each iteration of such a feedback loop must take at least 5 msec. Furthermore, if the firing rate was more typical (for example, 50 spikes per second), each loop would take 20 msec. Given that a great deal of visual processing can be accomplished in 100 msec, and a minimum of 10 synaptic stages are involved, the opportunity for the extensive use of feedback loops must be severely limited. However, one way in which feedback might be used is on a purely local basis. In the retina, for example, it is known that neuronal communication is achieved without the use of action potentials - signals are transmitted by the use of graded potentials. Feedback loops between adjacent neurons could be achieved using such graded potentials without suffering from the temporal constraints involved in spike generation (Koch et al, 1986). However, such a mechanism would clearly be restricted to a very local scale and could not be implemented in the case of geniculo-cortico-geniculate loops or for feedback between different cortical areas.

Relaxation. A number of connectionist models have proposed that visual processing involves a process related to relaxation, in which the system settles down to a minimum energy state compatible with the input configuration. Often, such models assume 50 or so cycles in which the activation levels of each processing element are allowed to influence each other. Without denying that such mechanisms may of considerable importance, particularly in longer timecourse cognitive processes, it would appear that normal visual processing has to be achieved with a smaller number of processing cycles.

Iteration. Many machine vision algorithms for contour finding, erosion, skeletonisation etc. rely on large numbers of iterations of particular algorithms. The temporal constraints on visual system functioning imply that if something equivalent occurs in the human brain, the number of iterations must be reduced to an absolute minimum, perhaps by the use of parallel processing algorithms. This is in itself an argument that machine visual algorithms must be "parallelisable".

Top-down influences and the role of context. It is widely assumed that one of the main reasons for why the human visual system is so much better than currently available machine vision systems is that humans can make use of context. That is to say, the knowledge that one is in a restaurant makes it easier to identify the pepper-mill on the table. This ability to use context is certainly important. However, the high level of performance of subjects in our visual identification task (see section 2.2), despite the fact that the subjects had no advance information about the type of image

with which they were to be presented indicates that even in the absence of contextual cues, the human visual system is capable of functioning remarkably well.

In this respect, it is important to realize that context can work in a variety of ways. Imagine an experiment in which you are presented with an unknown picture which you have to identify. Let us suppose that the picture is of a car on a motorway. The motorway provides a "within-picture" context which may help with the identification of the car. However, the fact that such contextual help is not essential is clear from the fact that you could also identify the same car against a plain white background. On the other hand, there is good evidence that an inappropriate context can severely impair visual identification (Palmer, 1975) - a drawing of a car in a living room is particularly difficult to identify. However, showing that an inappropriate context can make identification difficult does not mean that context is necessary for visual processing. Nevertheless, there is a clear distinction between this type of "within-picture" context, and the sort of context that would be provided if the picture of a car came in a series of pictures of different means of transport (train, aeroplane, bicycle etc.). This sort of continuous updating of the context on the basis of previous stimuli may well be important, but is not relevant to the question of the flow of information which occurs during the processing of a particular image by the visual system.

Temporalcoding.A final implication of these temporal constraints on visual processing concerns the possibility of various types of temporal coding schemes in the visual system. While traditional views of neural coding have generally assumed that information is primarily encoded by the firing rate of neurons, a number of authors have recently proposed that information is encoded in the temporal sequence of activity of individual neurons. For example, Lestienne and Strehler (1988, Strehler and Lestienne, 1986) have proposed that the temporal organization of triplets of spikes in individual neurons may itself carry information important for visual processing. Thus if a neuron emits three spikes with interspike intervals of 10 msec and 20 msec, this could have a different "meaning" to three spikes generated with intervals of 20 msec and 10 msec. One problem with such a model is that a delay of 30 msec would be introduced before a neuron at the next stage would be able to decode such information. Given that at minimum of 10 synaptic stages have to be negotiated in 100 msec, it is clear that there is insufficient time to use such a code at each stage, although it is perhaps possible that one or two processing stages might use such information.

Similar arguments apply to the model proposed by Richmond and coworkers (Richmond et al, 1987a,b,c) who proposed that the temporal sequence of activity in in inferotemporal cortex neurons contains information about the stimulus. While this may well be the case, their suggestion that changes in activity over a period of 100 to

200 msec could be used to code information could certainly not be used at each stage of visual processing - if each of 10 processing stages had to wait for 100 msec to process the information encoded by the neurons of the previous level, visual processing would take more than 1 second.

Finally, von der Mahlsburg and Bienenstock (1986, Bienenstock and von der Mahlburg, 1987) have suggested that information is carried in the degree of correlation between neurons. They estimate that 100 or 200 msec are needed for the degree of correlation between two neurons to be determined. While such a mechanism could be used in a simple 2 layer network composed of a "retinal" input layer and a "cortical" output layer, the fact that visual processing involves a minimum of 10 processing stages makes it difficult to see how such a mechanisms could allow the activation of highly selective temporal lobe neurons with such short latencies.

Thus it would appear that this examination of the anatomical, neurophysiological and psychological constraints on visual processing has wide ranging implications for the way information is processed by the visual system. In the next two sections we would like to deal with two specific points. The first concerns alternative ways of coding analog values which avoid the serious temporal constraints that arise if firing rate is used as a code. The second concerns possible alternative roles for the anatomically widespread feedback connections found in the visual system.

2.10 Alternative ways of coding analog information.

In the last section it was argued that the temporal constraints associated with visual information processing in humans make it unlikely that the firing rate of neurons could be used to code analog values. While it is true that in principle, neuronal firing rate can be used to code analog values with arbitrary levels of precision, the fact that in the visual system each processing stage has only about 10 msec of computation time means that a great deal of processing has to be done on the basis of only one spike in each of the neurons of the preceding layer. Nevertheless, there are other ways of increasing analog precision. One way is to take advantage of the massively parallel nature of visual processing. Imagine a neuron which receives inputs from 100 different neurons at the preceding stage, and that the probability that neurons in the preceding layer fire in response to a stimulus varies with stimulus intensity. Clearly the proportion of the 100 input neurons which are active will be a monotonic function of stimulus intensity - and this could be used to code as many as 100 different analog values. Such a scheme seems quite plausible in real brains but has the fairly obvious disadvantage of using large numbers of units.

However, there is another way of encoding analog information which may be of great importance. It is known that the time taken for a neuron to reach its firing threshold depends on the intensity of the stimulus. Thus a neuron which responds with a latency of 4 msec to an intense stimulus might take 10 msec or more in the case of a weak stimulus. Such properties are well known phenomena that result from the temporal summation properties of neurons. Imagine that an image is presented at time T_O. Certain parts of the image may contain contours with high contrast, whereas in other regions the contrast may be low. As a result, the first spikes to leave the retina would be those in the high contrast regions, whereas the first spikes would have longer latencies in the low contrast regions.

Thus information is presumably contained in the relative arrival times of action potentials coming from different sources. How might higher level neurons make use of such information? It is known that neurons in the auditory system are sensitive to differences in the arrival time of auditory stimuli at the two ears of around 1 msec (Knudsen et al, 1987), and recent work on the electric fish has shown that neurons can be sensitive to information in the microsecond range (Kawasaki et al, 1988). Poggio and Koch (1986) have proposed that sophisticated processing may occur in the dendritic trees of neurons. For example, imagine a dendrite on which two synapses make contact, an inhibitory terminal and an excitatory one, and that the inhibitory terminal is the closest to the cell body. If the excitatory signal arrives first, it would propagate along the dendrite without difficulty. However, if the inhibitory input arrives first, it could "veto" the passage of the second excitatory input. As a result, the cell itself could be sensitive to the relative arrival times of the two inputs,

responding only when they arrive in a particular order. In Poggio and Koch's description, such synaptic "veto" mechanisms were suggested to be involved in the calculation of stimulus motion. However, it seems reasonable to suggest that equivalent circuits could be used to pick up on analog information contained in spike arrival times.

Furthermore, spike arrival times do not have to be thought of as only dependent on such attributes as stimulus contrast and intensity. More generally, firing rate has often been thought of as a code for the probability that a particular feature is present (Barlow, 1972). The same could be true in the case of spike arrival times. If we consider a neuron in V2 receiving information from a variety of orientation-tuned neurons in V1, one can imagine that if spikes from the vertically-tuned neurons arrived first, this could provide additional information about stimulus orientation.

2.11 What do feedback circuits do?

In section 2.10 it was argued that a great of visual information processing must be achieved without the need for feedback pathways. However, this fails to account for one very obvious fact, namely that such pathways do exist, and are even numerically very important. In this section, we would like put forward a few suggestions of the other possible roles for feedback pathways.

Context. Although in the last section it was argued that the visual system can process visual information even in the absence of context, there is no doubt that contextual information can influence visual information processing. A large number of experiments have investigated "priming" effects in vision. For example, it is well established that the time taken by subjects to judge whether a presented letter string is a word or not (i.e. a lexical decision task) can be reduced if another semantically related word is presented just beforehand. Such effects may be due to top-down preactivation of the appropriate visual processing elements.

I**magery.** A possibly related process is that of visual imagery. Recent experiments have shown that visual processing in the occipital lobe can be influenced if the subject imagines the presence of a stimulus in a particular part of the visual field (Farah et al, 1988). It is possible that such effects are mediated by descending pathways to the visual cortex from higher order areas such as the temporal lobe.

Attention. When a subject focuses his attention on a particular part of the visual field, his ability to detect a stimulus presented in that region is improved (Posner, and Presti, 1987). This attention-related sensory enhancement is quite possibly related to the enhancement of the responsiveness of single neurons in awake primates described by Wurtz and his collaborators (Wurtz, Goldberg and Robinson, 1980). Such excitability changes could well be mediated by descending influences from higher order areas.

Ambiguous stimuli. There are numerous situations in which the interpretation of a particular visual pattern is ambiguous. A classic case is the Necker cube, which can be seen in one of two different configurations. Such ambiguous figures can often be flipped from one state to the other under voluntary control - perhaps this could reflect another case of top-down influences on perception.

Conflicting stimuli. If different images are presented to the two eyes which cannot be fused, one typically experiences a phenomenon known as binocular rivalry. For example, if the left eye sees the world through a red filter and the right eye through a green one, each particular part of the visual field is either seen as red or green, but not both. It is possible that corticogeniculate pathways may be involved in suppressing one or other eye's input in such rivalrous situations.

Learning. Considerable attention has recently been paid to the possibility of learning in neural networks using the Back-propagation learning algorithm (Rumelhart, Hinton and Williams, 1986). Although as we will see in a later section, there is some doubt as to whether such an algorithm is implemented in real brains, if it were to be implemented, then feedback pathways could be involved. The algorithm requires that an error signal from output neurons be back-propagated to neurons in preceding levels. In principle, this could be achieved by either sending signals backwards along the originally activated axons, or alternatively, by using topographically precise reverse projections from the output layer to units in the hidden layers.

Thus, even if it is the case that in normal visual processing, feedback loops are only used sparingly, there is no shortage of candidate roles for the large number of feedback pathways which have been found anatomically.

2.12 Processing - Final conclusions.

Current connectionist models are only fairly loosely constrained by biological data. However, if we restrict ourselves to a particular domain, namely visual information processing, biological constraints can be used to specify in considerably greater detail the main features of a connectionist visual system designed to simulate the capacities of the human visual system. This constraints can be summarized as follows.

A. The visual system is arranged as a massively parallel multilayer feedforward net with at least 10 processing layers.

B. Considerable visual analysis is possible with a single forward pass through the network.

C. In many situations, each unit can only emit 1 spike before the units in the next layer have to respond.

D. Firing rate per se cannot be used during visual processing to code analog values with any real precision.

E. Coding of analog values could however be achieved by making use of the arrival times of spikes from different sources - the earliest arriving signals could be given priority.

F. Sophisticated dendritic processing could mean that each unit could be doing more than simply calculating the sum of all the input activations - logical "and", "and not" functions could well make the system highly non-linear.

G. Although feedback pathways between different layers are present, there may not be time to use them during normal visual processing. They could however play a role in the effects of context, imagery, attention, resolution of ambiguous stimuli and learning.

H. The use of iterative loops is kept to an absolute minimum and perhaps even eliminated by the use of massive parallelism.

3. LEARNING

3.1 Introduction - Different learning algorithms.

One of the main reasons for the current explosion of interest in connectionist modelling is that such systems can self-organize. This ability depends on the implementation within the network of some sort of learning algorithm. Currently, there are a number of learning rules available, including variations of the Hebb learning rule, the Perceptron learning rule, Back-propagation and the Boltzmann learning algorithm. (For reviews, see Sejnowski, 1988; Arbib, 1988; Soucek and Soucek, 1988).

As was the case with processing, one can take one of two positions. One possibility is to say that we do not mind whether the algorithm we use is biologically realistic - the only important criterion is whether it works or not. Alternatively we can take the view that it is important to try understand the only truly intelligent system we have that we are sure can actually work - namely, real biological brains.

In this section, we would like to look at some of the evidence relating to the biological plausibility of some of these different rules. It will be argued that whereas there is considerable reason to believe that something like Hebbian learning occurs in the brain, there is currently little clear evidence for certain other learning mechanisms, and in particular Back-propagation.

3.2 Hebbian learning rules.

In his classic book, "The Organization of Behavior", published in 1949, the Canadian psychologist Donald Hebb proposed the following neurophysiolgical postulate :

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased" (page 62)

This suggestion, which has since become known as Hebb's postulate, has been widely used in theoretical analyses of neural function. In fact, several variants have been suggested, but they all have the basic feature that the connection strength between two neurons depends on the degree of coactivation of the two neurons. In the most elementary form there is an increase in connectivity if a synaptic terminal and its post-synaptic target cell are active simultaneously. There is now a considerable amount of neurophysiological evidence favouring the existence of such synaptic modification mechanisms. For example, Frégnac, Shultz, Thorpe and Bienenstock (1988) were able to modify the response properties of neurons in the cat visual cortex using an experimental protocol which artificially changed the degree of coactivation between pre- and post-synaptic elements. Visual cortical neurons typically   show orientation selectivity, responding best to certain orientations. In experiments, Frégnac and coworkers first determined the orientation selectivity of a cortical neuron, and then used a conditioning procedure in which two orientations were repeatedly presented - the preferred orientation, and a second orientation to which the neuron responded, but not as strongly. During this conditioning, every time the preferred orientation was presented, the activity of the neuron was artificially suppressed by passing current through the recording electrode. By contrast, every time the non-preferred orientation was presented, the neurons activity was artificially increased. Subsequently, it was found that many neurons changed their   orientation preference in favour of the orientation associated with high levels of neuronal activity. Thus, such evidence supports the basic premise of Hebb's postulate, namely the idea that synapses that are active at the same time as the post-synaptic neuron can become strengthened. Other support can be found in the experiments of Baranyi and Feher (1981) who were able to increase the strength of synaptic inputs to cat motor cortex neurons by artificially activating the neuron every time the synapse was active, and in a number of studies of synaptic plasticity in the hippocampus (e.g. Kelso et al, 1986).

These experiments support the original proposition made by Hebb, namely that synaptic efficiency increases when the pre and post-synaptic elements are active at the same time. However, there are other variations on the basic Hebb rule.   One possibility is that synapses which are inactive when the cell is active become less

effective in driving the cell (see Sejnowski, 1977), and Singer (1987) has recently argued for such a mechanism.

A fundamentally different suggestion is to propose that synaptic modification can occur even when the post-synaptic element is inactive. Until recently, the idea that changes can occur in the absence of neuronal activity has been only a theoretical possibility. However, Reiter and Stryker (1988) have provided experimental evidence which implies that such a process does indeed occur. It has been known since the early 1960s that if a kitten is subjected to monocular deprivation, the eye that has normal visual experience comes to dominate visually responsive cells in the visual cortex. This phenomenon can be readily explained on the basis of Hebb-like mechanisms - only synapses driven by the exposed eye can be reinforced because only they can be activated visually. Reiter and Stryker used a variant on this technique in which during the period of monocular deprivation, the activity of visual cortical neurons was locally suppressed by directly infusing an inhibitory substance into the cortex. Subsequently, they found that the eye that had been occluded was actually better at driving the cortical cells than the eye which had received normal visual experience. This seemingly counter-intuitive result is actually consistent with a particular type of coactivation based modifiability - when there is no post-synaptic activity, inactive synapses have an advantage, but as yet it is unclear whether this is because the inactive synapses become more effective, or whether the active synapses become less effective.

Thus, there is now a considerable body of experimental data in support of the idea that the degree of coactivation between pre and post-synaptic elements can lead to modifications in synaptic weight. Such correlation based learning rules have been widely used in a variety of theoretical studies, including matrix memory models (Amari, 1977; Kohonen, 1977, 1984; Palm, 1982), reinforcement learning (Sutton and Barto, 1981), competitive learning (Fukushima, 1975; Fukushima and Miyaki, 1982; Grossberg, 1982; Rumelhart and Zipser, 1985, von der Mahlsburg, 1973), the formation of higher level response properties in multilayer networks (Linsker, 1986a,b,c) and Hopfield nets (Hopfield, 1982).

3.3 Error correction rules.

Another common type of learning algorithm uses an error signal to modify synaptic weights. This involves a calculation of the difference between the desired output of a unit and its actual state. The resulting error signal is then used to control the amount of synaptic weight change. For example, in the Perceptron model developed by Frank Rosenblatt and his colleagues (Rosenblatt, 1962), if an output unit is active when it should have been inactive, active inputs are decreased in weight, while inactive inputs are increased in weight. Conversely, if the output unit is inactive

when it should have been active, active synaptic inputs are increased in weight whilst those which are inactive become less effective. The net effect is to minimize errors. Similar mechanisms underlie the Widrow-Hoff rule, and the recently developed Back-propagation algorithm (although in the latter case, the error signals in the output units can be used to adjust weights on intermediate "hidden" units).

While there is no doubt that such error-correction rules have considerable power, the question needs to raised as to whether or not such mechanisms actually occur in real nervous systems, and in particular in the human brain. As far as we are aware there is as yet no direct neurophysiological evidence for such mechanisms, although it has to be said that there have probably not been any experiments specifically designed to test such a notion.

Nevertheless, there are certain theoretical arguments that indicate that at least compared with the coactivation related synaptic learning rules discussed in the previous section, error correction mechanisms may be relatively difficult to implement in real neurons. There is now good reason to think that coactivation based modification can be achieved on a purely local basis, and indeed considerable progress has recently been made on the cellular processes that could underlie such modifiability (Artola and Singer, 1987, Cotman et al, 1988). When neural networks with error correction are simulated on a computer, the question of how to calculate the error signal is minimized - you just subtract one floating point value from another. However, in a real neural net the problem is not so trivial and may well require a minimum circuit of three neurons - one neuron would signal the desired activity level, another would signal the actual activity level generated by the network, and a third neuron would be required to signal the difference between these two levels of activation. While potentially the anatomical substrate for such circuits exists in the brain, it certainly involves a greater level of complexity compared with local coactivation based modification.

3.4 Back-propagation.

One particular kind of error-correction rule has received a great deal of attention in the last few years. This is the Back-propagation algorithm discussed in detail by Rumelhart, Hinton and Williams (1986). This algorithm can be used to modify weights in multi-layer networks, something which earlier error-correction rules like the Perceptron learning rule were unable to do. Indeed, this gets round one of the major criticisms made by Marvin Minsky and Seymour Papert in their classic book on Perceptrons published in 1969. They proved that a single layer Perceptron was unable to solve certain types of pattern classification problem, such as the parity problem (respond if an odd number of inputs are active), and problems involving "exclusive-or" (XOR) rules (respond if one or other but not both of a pair of inputs are

active). At the time it was felt that such limitations would prevent neural networks forming the basis of a truly intelligent artificial system, and this led to the great majority of researchers in artificial intelligence turning to symbol manipulation paradigms during the 1970s. The development of the back-propagation algorithm has changed all that, because now it is possible to train multi-layer Perceptron like networks to solve previously intractable pattern classification problems, and in the last few years we have witnessed a veritable explosion of interest in neural networks.

However, is it really the case that an intelligent system must be able to train layers of hidden units in order to be worthy of the label "intelligent". What evidence is there for this type of learning in humans? Recently, Thorpe, O'Regan and Pouget (1988) tried to examine this question by looking at the ability of humans to solve pattern classification problems of the type that single layer Perceptrons are unable to solve. The subject was presented with a series of random configurations of high or low intensity squares arranged in a 4 by 4 array on the screen of a computer terminal. At the center of the array was a fixation point, which the subject was instructed to fixate. New configurations were presented at a rate of 5 per second, accompanied by a beep if the pattern fulfilled some particular rule. Various boolean rules were used, including "simple" rules (beep if a particular square was on), "and" rules (beep if particular pair of elements are both on) and "Xor" rules (beep if one or other of a pair of squares is on, but not both). Varying numbers of training patterns were presented (10, 20, 30, 40, 60, 100, 160 or more) after which the subject had to indicate which of the 16 squares were relevant. For each type of problem ("simple", "and" and "Xor") and for each training sequence length, the entire procedure was repeated 100 times in order to calculate the percentage of runs on which the subject correctly identified the relevant squares.

It was found that "simple" and "and" rules could be solved rapidly - often the solution could be found after only 40-60 training patterns. However, in the case of "XOR" rules, there was no evidence for learning even after 4000 training examples had been presented. By contrast, a three-layer neural network using back-propagation, while relatively slow at solving equivalent "simple" and "and" problems, was consistently able to solve the problems involving "Xor" relationships. Although it would certainly be premature to conclude that humans are unable to solve any problems involving "XOR" relations, it would appear that at least in this particular case, there is as yet no reason to believe that they can. What are the implications of such findings?

The first point would seem to be that the ability to solve problems such as Xor, parity and so on should not necessarily be thought of as the hallmark of an intelligent system. Human beings evolved cognitive capacities that were useful to them in the evolutionary struggle for survival. For this, one can well imagine that the ability to rapidly spot correlations between events in the outside world (i.e. conjunctions) could

be of great importance. It is therefore perhaps not so surprising that humans seem to be remarkably good at detecting correlations between the presence of a particular square in a visual array and the presentation of a beep. By contrast, it is really quite difficult to think of a natural situation in which the ability to solve an exclusive or problem would confer and advantage. In general, if two cues both signal that food is about to arrive, when the two are present at the same time, the food is even more likely to appear!

Even if natural events involving Xor relations do occur, one could well argue that if at least 2000 training sessions are required in order to find the underlying rule (as was the case for the network in our simulation), the advantage for an individual who possessed the means for solving such problems relative to his less sophisticated competitor would be only marginal - how often is he going to get 2000 chances to observe the relevant events? If, however, the number of relevant features is small, and new patterns are presented only relatively infrequently, then it may be possible to solve the problem by sequential reasoning - but this is hardly what is thought to occur in a back-propagation programmed neural network.

But there is a second reason to think that an ability to solve Xor type problems would be difficult to evolve. Currently, the back-propagation algorithm would appear to be the best tool we have for solving Xor problems with neuron-like networks. Is back-propagation feasible neurophysiologically? In the previous section we have already pointed out that the calculation of error signals in real nervous systems may not be as easy as it appears when doing neural network simulations on a digital computer. However, in addition to this, back-propagation requires that this error signal be sent back to neurons at lower levels in the processing sequence. With a computer simulation, it is trivial to "propagate" the resulting error value to the input units - but how could this be done in a real nervous system? Retrograde transport of chemical substances along axons is known to exist but it is generally a rather slow process and would be unlikely to have the temporal resolution necessary for processing the 2000 training trials that would be necessary. Another possibility, namely that the error signal could be sent along separate feedback pathways, would require a topographic precision that would be difficult to imagine in a real nervous system. All of these considerations make the evolution of a new neurophysiological mechanism capable of supporting the back-propagation algorithm unlikely.

3.5 Learning - Final Conclusions.

We have argued that whereas in the case of Hebb-like coactivation based synaptic weight changes there is now reasonably good experimental evidence that such mechanisms actually occur in real brains, there is as yet little evidence that either error-correction learning rules or back-propagation are implemented in

biological neural networks.   However, it is important to stress that even if the human brain turns out to be a highly intelligent system that manages without back-propagation, this should not be taken as an argument for reducing the current intense research effort into the development of such tools. There is no reason why researchers working on artificial neuron-like systems should be subject to the same constraints that have applied during the evolution of the human nervous system. Firstly, in an artificial nervous system, the calculation of error and its propagation pose no particular problems. Secondly, whereas under natural conditions, animals are unlikely to experience the several thousands of training trials necessary to resolve complex problems, the presentation of large numbers of examples to modern digital computers can be done very rapidly.

On the other hand, if humans do achieve a high level of intelligence without using back-propagation, this could be seen as an argument for investing a significant research effort into simpler mechanisms. Coactivation-based synaptic modification has been used in a large number of theoretical studies, and it seems unlikely that we have done more than scratch the surface of the full potential of such mechanisms for the development of intelligent artificial systems.

4. PERSPECTIVES

In the introduction, we pointed out that connectionist modelling can be used in two fairly distinct ways. One approach is aimed at the development of sophisticated artificial intelligence, the other is aimed at understanding how real brains work, and in particular the mechanisms underlying the functioning of the human brain. It was suggested that computer scientists and engineers endeavoring to design and construct artificial systems could in principle use "neurally-inspired" connectionist architectures without worrying too much about whether or not they are biologically realistic. While this is true, we feel strongly that greater information flow between researchers in these different fields will be of great benefit. In this article, we have reviewed two particular areas of research, namely information processing by the human and primate visual system, and learning algorithms in biological systems. There can be little doubt that the performance of the human brain on tasks such as visual object identification and pattern classification is highly impressive. And yet the evidence suggests that the underlying mechanisms may not be completely intractable.

The human brain has the great advantage of being a compact connectionist system which really works - all of us have in our possession a highly perfected system which has had virtually all the "bugs" removed by millions of years of natural selection. The processing architectures and learning algorithms have been thoroughly tried and tested. Connectionist modelling, lying as it does between the natural and the artificial, is the perfect domain for demonstrating that genuine interdisciplinary co-operation can reap enormous benefits.

5. REFERENCES

Amari, S. (1977) *Neural theory of association and concept formation.* Biological Cybernetics, **26**, 175-185.

Arbib, M.A. (1988) *Brains, machines and mathematics, 2nd edition.* New York : Springer-Verlag.

Artola, A. and Singer, W. (1987) *Long-term potentiation and NMDA receptors in the visual cortex.* Nature, **330**, 649.

Baranyi A. and Féher O. (1981) Nature, **290**, 413-415.

Barlow, H.B. (1972) *Single units and sensation : A neuron doctrine for perceptual psychology.* Perception, **1**, 371-394.

Bienenstock E. and von der Mahlsburg C. (1987) *A neural network for invariant pattern recognition.* Europhysics Letters, **4**, 121-126.

Biederman, I. (1987) *Recognition by components : A theory of human image understanding.* Psychological Review, **94**, 115-147.

Bullier, J., McCourt, M.E. and Henry, G.H. (1988) *Physiological studies on the feedback connection to the striate cortex from cortical areas 18 and 19 in the cat.* Experimental Brain Research, **70**, 90-98.

Carr, T.H., McCauley, C., Sperber, R.D., and Parmelee, C.M. (1982) *Words, pictures and priming : On semantic activation, conscious identification and the automaticity of information processing.* Journal of Experimental Psychology : Human Perception and Performance, **8**, 757-777.

Cattell, J.M. (1886a) *The time it takes to see and name objects.* Mind, **11**, 63-65.

Cattell, J.M. (1886b) *The time taken up by cerebral operations.* Mind, **11**, 277-292.

Cotman, C.W., Monaghan, D.T. and Ganong, A.H. (1988) *Excitatory amino acid neurotransmission : NMDA receptors and Hebb-type synaptic plasticity.* Annual Review of Neuroscience, **11**, 61-80.

Dawkins, R. (1986) *The blind watch-maker.* London : Longmans.

DeYoe, E.A. and Van Essen, D.C. (1988) *Concurrent processing streams in monkey visual cortex.* Trends in Neurosciences, **11**, 219.

Dinse, H. and Best, J. (1988) Paper presented at the First European Conference on Neural Networks, Paris.

Farah, M.J., Péronnet, F., Gonon, M.A. and Giard, M.H. (1988) *Electrophysiological evidence for a shared representation medium for visual images and visual percepts.* Journal of Experimental Psychology : General, **117**, 248-257.

Feldman, J.A. (1985a) *Connectionist models and parallelism in high level vision.* In A. Rosenfeld (ed.), *Human and Machine Vision II.* (pp 86-108). Orlando : Academic Press.

Feldman, J.A. (1985b) *Four frames suffice : A provisional model of vision and space.* Behavioral and Brain Sciences, **8**, 265-289.

Feldman, J.A. and Ballard, D.H. (1982) *Connectionist models and their properties.* Cognitive Science, **6**, 205-254.

Frégnac, Y., Shultz, D., Thorpe, S. and Bienenstock, E. (1988) *A cellular analog of visual cortical plasticity.* Nature, **333**, 367-370.

Fukushima, K. (1975) *Cognitron : A self-organizing multilayered neural network.* Biological Cybernetics, **20**, 121-136.

Fukushima, K. and Miyaki, S. (1982) *Neocognitron : A new algorithm for pattern recognition tolerant of deformations and shifts in position.* Pattern Recognition, **15**, 455-469.

Grossberg, S. (1982) *Studies of Mind and Brain.* Reidel.

Hebb D.O. (1949) *Organization of Behavior.* New York : Wiley.

Henderson, J.M., Pollatsek, A. and Rayner, K. (1987) *Effects of foveal priming and extra-foveal preview on object identification.* Journal of Experimental Psychology : Human Perception and Performance, **13**, 449-463.

Heywood, C.A. and Cowey, A. (1988) *On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys.* Journal of Neuroscience, **7**, 2601-2617.

Hildreth, E.C. and Koch, C. (1987) *The analysis of visual motion : From computational theory to neuronal mechanisms.* Annual Review of Neuroscience, **10**, 477-534.

Hopfield J. (1982) *Neural networks and physical systems with emergent collective computational properties.* Proceedings of the National Academy of Science, **79**, 2554-2558.

Hinton, G.E. and Anderson, J.A. (1981) *Parallel models of associative memory.* Hillsdale : Lawrence Erlbaum.

Humphreys, G.W. and Quinlan, P.T. (1988) *Normal and pathological processes in visual object constancy. I*n G.W. Humphreys and M.J. Riddoch (eds) V*isual object recognition : A cognitive neuropsychological approach.* (pp. 43-105) Hillsdale : Lawrence Erlbaum.

Imbert, M. (1988) *La vision naturalle : Le traitement neuronal de l'information visuelle.* Intellectica, **5**, 3-31.

Intraub, H. (1979) *The role of explicit naming in pictorial encoding.* Journal of Experimental Psychology : Human Learning and Memory, **5**, 78-87.

Intraub, H. (1980) *Presentation rate and the representation of briefly glimpsed pictures in memory.* Journal of Experimental Psychology : Human Learning and Memory, **6**, 604-610.

Intraub, H. (1981) *Rapid conceptual identification of sequential presented pictures.* Journal of Experimental Psychology : Human Perception and Performance, **7**, 604-610.

Jeffreys, A. (1988) *A human scalp-potential sensitive to faces.* Perception, **17**, A31.

Kawasaki, M., Rose, G. and Heiligenberg, W. (1988) *Temporal hyperacuity in single neurons of electric fish.* Nature, **336**, 173-176.

Kelso, S.R., Ganong, A.H. and Brown, T.J. (1986) *Hebbian synapses in hippocampus,* Proceedings of the National Academy of Science, **83**, 5326-5330.

Kendrick, K.M. and Baldwin, B.A. (1986) *The activity of neurons in the lateral hypothalamus and zona incerta of the sheep responding to the sight or approach of food is modified by learning and reflects food preference.* Brain Research, **375**, 320-328.

Kendrick, K.M. and Baldwin, B.A. (1987) *Cells in the temporal lobe of the sheep can respond preferentially to the sight of faces.* Science, **236**, 448.

Knudsen, E.I., du Lac, S. and Esterley, S.D. (1987) *Computational maps in the brain.* Annual Review of Neuroscience, **10**, 41-66.

Koch, C., Marroquin, J. and Yuille, A. (1986) *Analog "neuronal" networks in early vision.* Proceedings of the National Academy of Science, **83**, 4263-4267.

Kohonen, T. (1977) *Associative memory : A system-theoretical approach.* Berlin : Springer-Verlag.

Kohonen, T. (1984) *Self-organization and associative memory.* Berlin: Springer-Verlag.

Lehky, S.R. and Sejnowski, T.J. (1988) *Network model of shape-from-shading : neural function arises from both receptive and projective fields.* Nature, **333**, 452-454.

Lestienne, R. and Strehler, B.L. (1988) *Differences between visual cortical cells in triplet and ghost doublet informational symbols relationships.* Biological Cybernetics, **59**, 337-352.

Linsker, R. (1986a) *From basic network principles to neural architecture : Emergence of spatial opponent cells.* Proceedings of the National Academy of Science, **83**, 7508-7512.

Linsker, R. (1986b) *From basic network principles to neural architecture : Emergence of orientation selective cells.* Proceedings of the National Academy of Science, **83**, 8390-8394.

Linsker, R. (1986a) *From basic network principles to neural architecture : Emergence of orientation columns.* Proceedings of the National Academy of Science, **83**, 8779-8783.

Lund, J.S. (1988) *Anatomical organization of macaque monkey striate cortex.* Annual Review of Neuroscience, **11**, 253-288.

Martin, K.A.C. (1988) *From enzymes to perception : A bridge too far?* Trends in Neurosciences, **11**, 380-387.

Maunsell, J.H.R. and Newsome, W.T. (1987) *Visual processing in monkey extrastriate cortex.* Annual Review of Neuroscience, **10**, 363-402.

Minsky, M. and Papert S. (1969) *Perceptrons.* MIT Press.

Minsky M. and Papert S. (1988) *Perceptrons. Expanded Edition.* MIT Press.

Mishkin, M., Ungerleider, L.G. and Macko, K.A. (1983) *Object vision and spatial vision : Two cortical pathways.* Trends in Neurosciences. **6**, 414-416.

Newsome, W.T. and Wurtz R.H. (1988) *Probing visual cortical function with discrete chemical lesions.* Trends in Neurosciences, **11**, 394-399.

Oldfield, R.C. (1966) *Things, words and the brain.* Quarterly Journal of Experimental Psychology, **18**, 340-353.

Oldfield, R.C. and Wingfield, A. (1965) *Response latencies in naming objects.* Quarterly Journal of Experimental Psychology, **17**, 273-281.

Palm, G. (1982) *Neural Assemblies.* Berlin : Springer-Verlag.

Palmer, S.E. (1975) *The effect of contextual scenes on the identification of objects.* Memory and Cognition, **3**, 519-526.

Perrett D.I., Rolls E.T. and Caan W.C. (1982) *Visual neurons responsive to faces in the monkey temporal cortex.* Experimental Brain Research, **47**, 329-342.

Perrett, D.I., Mistlin, A.J. and Chitty A.J. (1987) *Visual neurons responsive to faces.* Trends in Neurosciences, **10**, 358.

Poggio, T. and Koch, C. (1986) *Synapses that compute motion.* Scientific American, **256**, 46-71.

Posner, M.I. and Presti, D.E. (1987) *Selective attention and cognitive control.* Trends in Neurosciences, **10**, 13.

Potter, M.C. (1975) *Meaning in visual search.* Science, **187**, 965-966.

Potter, M.C. (1976) *Short-term conceptual memory for pictures.* Journal of Experimental Psychology : Human Learning and Memory, **2**, 509-522.

Potter, M.C. and Faulconer, B.A. (1975) *Time to understand pictures and words.* Nature, **253**, 437-438.

Reiter, H.O. and Stryker, M.P. (1988) *Neural plasticity without post-synaptic action potentials : Less active inputs become dominant when kitten visual cortical cells are pharmacologically inhibited.* Proceedings of the National Academy of Science, **85**, 3623-3627.

Richmond, B.J. Optican, L.M. Podell, M, and Spitzer H. (1987a) *Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics.* Journal of Neurophysiology, **57**, 132-146.

Richmond, B.J. and Optican, L.M. (1987b) *Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform.* Journal of Neurophysiology, **57**, 147-161.

Richmond, B.J. and Sato, T. (1987c) *Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis..* Journal of Neurophysiology, **57**, 161-178.

Riddoch, M.J. and Humphreys, G.W. (1988) *Picture naming.* In G.W. Humphreys and M.J. Riddoch (eds.) *Visual Object Processing : A Cognitive Neuropsychological approach.* (pp107-143).

Rolls, E.T. (1986) *Information representation, processing and storage in the brain : analysis at the single unit level.* In J.P. Changeux and M. Konishi (eds), *Neural and molecular mechanisms of learning.* Berlin : Springer-Verlag.

Rolls E.T., Burton M. and Mora F. (1976) *Hypothalamic neuronal responses associated with the sight of food.* Brain Research, **111**, 53-66.

Rolls E.T., Perrett D.I., Caan W.C. and Wilson F.A.W. (1982) *Neuronal responses related to visual recognition.* Brain, **105**, 611-646.

Rolls E.T., Sanghera M. and Roper-Hall A. (1979) *The latency of activation of neurones in the lateral hypothalamus and substantia innominata during feeding in the monkey.* Brain Research, **164**, 121-135.

Rosenblatt F. (1962) *Principles of Neurodynamics.* Spartan Books.

Rumelhart D., Hinton G. and Williams D. (1986) *Learning internal representations by error propagations.* In Rumelhart and McClelland (eds.)Vol 1. (pp 318-362).

Rumelhart, D.E. and McClelland J. (1986) *Parallel Distributed Processing : Explorations in the microstructure of cognition.* MIT Press/ Bradford Books.

Rumelhart, D.E. and Zipser D. (1985) *Feature discovery by competitive learning.* Cognitive Science, **9**, 75-112.

Sejnowski, T.J. (1977) *Storing covariance with nonlinearly interacting neurons.* Journal of Mathematical Biology, **4**, 303-321.

Sejnowski, T.J. (1988) *Neural network learning algorithms.* In R. Eckmiller and C. von der Mahlsburg (eds.) *Neural computing.* (pp 291-300).

Sejnowski, T.J., Koch, C. and Churchland, P.S. (1988) *Computational Neuroscience,* Science, **241**, 1299-1306.

Shevelev, I.A. (1987) *Temporal processing of orientation signals in the visual cortex.* Perception, **16**, 263-264.

Singer W. (1987) *Activity dependent self-organization of synaptic connections as a substrate of learning.* In J.P. Changeux and M. Konishi (eds.) *The neural and molecular bases of learning.* (pp. 301-336).

Smolensky, P. (1988) *On the proper treatment of connectionism.* Behavioural and Brain Sciences, **11**, 1-76.

Soucek, B. and Soucek, M. (1988) *Neural and massively parallel computers.* New York : Wiley.

Strehler, B.L. and Lestienne, R. (1986) *Evidence on precise time-coded symbols and memory of patterns in monkey cortical neuronal spike trains.* Proceedings of the National Academy of Science, **83**, 9812-9816.

Sutton, R.S. and Barto, A.G. (1981) *Toward a modern theory of adaptive networks : Expectation and prediction.* Psychological Review, **88**, 135-170.

Thorpe, S.J. (1988) *Identification of rapidly presented images by the human visual system.* Perception, **17**, A77.

Thorpe, S.J. and Pouget, A. (1989) *Coding of orientation by the visual cortex : Neural network modelling.* In "Connectionism in Perspective".

Thorpe, S.J., O'Regan, J.K and Pouget, A. (1988) *Humans fail on XOR pattern recognition problems.* Proceedings of the First European conference on Neural Networks.

Thorpe, S.J., Rolls, E.T. and Maddison, S. (1983) *The orbitofrontal cortex : Neuronal activity in the behaving monkey.* Experimental Brain Research, **49**, 93-115.

Van Essen, D.C. and Maunsell, J.H.R. (1983) *Hierachical organization and functional streams in visual cortex.* Trends in Neurosciences, **6**, 370-375.

von der Mahlsburg, C. (1973) *Self-organizing of orientation sensitive cells in the striate cortex.* Kybernetik, **14**, 85-100.

von der Mahlsburg, C. and Bienenstock, E. (1986) *A neural network for the retrieval of superimposed connection patterns.* In F. Fogelman, F. Weisbuch and E. Bienenstock (eds.) *Disordered systems and biological organization.* Berlin : Springer-Verlag.

Wild, H.M, Butler, S.R., Carden, D., and Kulikowski, J.J. (1986) *Primate cortical area V4 important for colour constancy but not wavelength discrimination.* Nature, **313**, 133-134.

Wurtz, R.H., Goldberg, M.E. and Robinson, D.L. (1980) *Behavioral modulation of visual responses in the monkey : stimulus selection for attention and movement.* Progress in Psychobiology and Physiological Psychology, **9**, 43-83.

Zeki, S. (1980) T*he representation of colour in the cerebral cortex.* Nature, **284,** 412-418.

Zeki S. and Shipp S. (1988) *The functional logic of cortical connections.* Nature, **335**, 311-316.

Zipser, D. and Anderson R.A. (1988) *A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons.* Nature, **331**, 679-684.