CHAPTER 14
A COMPARATIVE ANALYSIS OF NEURAL MECHANISMS, RECENT DATA, AND ALTERNATIVE MODELS

Stephen Grossberg

14.1. Comparative Analysis of Neural Models

The previous chapters illustrate how contemporary neural network models provide insights into some of the organizational principles that govern biological sensory-motor systems, and offer a level of computational precision that enables sharp comparisons and contrasts to be made between different sensory-motor systems. The capacity of these models to clarify, integrate, and predict behavioral and neural data is predicated upon the coordinated use of theoretical, mathematical, computational and empirical tools in a manner that reveals many more constraints on brain design than empirical tools alone. No single experimental paradigm in the behavioral and brain sciences provides sufficiently many data to uniquely characterize a neural system. Interdisciplinary theoretical and empirical approaches that can coordinate and discover both top-down and bottom-up constraints at multiple levels of behavioral and neural organization provide a much greater level of guidance towards characterizing brain designs.

Where insufficient guidance to derive a unique model is provided by all the available top-down and bottom-up theoretical and data constraints, it is advisable to develop fully the major model types that are consistent with all known constraints. Such a comparative approach to model development was followed, for example, in Chapters 3, 5, 6, 10, and 11. By rendering explicit the theoretical relationships among each model's parts, a relatively small number of new experiments can efficiently rule out many otherwise plausible solutions to a prescribed neural problem. A comparative approach to neural modelling is also advisable to cope with the facts of evolutionary variation. One model realization of a fundamental neural design may exist in some species, whereas a different realization may have evolved in other species. One helpful way to understand the invariant properties of a neural design is to carefully articulate its major design variations. Last but not least, a comparative analysis also provides neural network technologists with several ways to implement a desired functional competence.

The present chapter carries out a comparative analysis in two different senses. First it compares variations on an important neural design that is realized differently within the eye and arm movement systems, as well as at different processing stages of the same system. This type of comparative analysis helps to sharpen our understanding of each system by clarifying why it is computationally different from other systems that share

share some of its design properties. Then comparisons of our models are made with subsequent models and data of other investigators. This type of comparative analysis may help the reader to form a more unified understanding of the complex and often fragmented literature on adaptive sensory-motor control.

14.2. Comparative Analysis of Movement Vectors in Eye and Arm Movements

The design principle of vector encoding has been described in both the control of saccadic eye movements by the superior colliculus (Mays and Sparks, 1980, 1981; Sparks, 1978; Sparks and Jay, 1987; Sparks and Mays, 1981) and the control of arm movements by the motor cortex (Evarts and Tanji, 1974; Georgopoulos, 1986; Georgopoulos, Kalaska, Caminiti, and Massey, 1982; Georgopoulos, Kalaska, Crutcher, Caminiti, and Massey, 1984; Georgopoulos, Kettner, and Schwartz, 1988; Georgopoulos, Schwartz, and Kettner, 1986; Kettner, Schwartz, and Georgopoulos, 1988; Schwartz, Kettner, and Georgopoulos, 1988; Tanji and Evarts, 1976).

Although neural vectors are encoded in both of these systems, they are encoded differently in each system. Why do both systems employ a neural code which, in its broad outline, seems similar, but in its computational realization is grossly different? Models of saccadic eye movements (Chapter 4) and of arm movements (Chapter 13) have been developed in which these distinct types of vector coding are utilized. The models suggest that the design constraints which force these differences concern different problems of learning and coordinate transformation that are solved by the eye and arm movement systems. These constraints were necessarily scattered across several of the previous chapters in order to develop each of them technically. The present discussion summarizes and compares these constraints in a single place, and suggests additional experiments that may be used to further test the models.

14.3. Map Vectors and Difference Vectors

Because different types of vectors seem to exist, we need simple but evocative names to clearly distinguish them. The experiments of Sparks and his colleagues on the deeper layers of monkey superior colliculus have revealed a type of vector coding that may conveniently be called a *vector map code* (see pp.307–308). In such a vector representation, each location in a spatially organized map encodes a different vector, hence, a different combination of movement length and direction. The most eccentric locations tend to code the longest movements. Changing the polar angle of a location tends to change movement direction. Exciting cells at a prescribed location in the map tends to cause a saccadic eye movement of corresponding length and direction. Populations of cells are activated under normal behavioral conditions. The eye movement triggered by such a population has a length and direction that corresponds to the average length and direction coded by all of the cells in the population.

The experiments of Georgopoulos and his colleagues on monkey motor cortex have revealed a type of vector encoding that may conveniently be called a *vector difference code* (see p. 316). In such a vector representation, each cell tends to generate a broad unimodal tuning curve of direction preference that may include 180° of movement directions. Coding of movement amplitude tends to covary with the firing rate of cells in their direction of maximal sensitivity. Moreover, changes in the initial position of the arm covary with changes in the baseline level of cell activity. Thus a vector difference code does not encode amplitude and direction via different locations in a spatial map.

The computation and design significance of these map vectors and difference vectors are clarified from the viewpoint of the system's total behavioral competence.

14.4. Vector Integration to Endpoint and GO Signal Modulation in Arm Movement Control

Difference vectors are found in the model of arm trajectory control, called the Vector Integration to Endpoint (VITE) model, that was described in Chapter 13. This model predicts how read-out of a target position command, or TPC, that specifies a desired target position, or motor expectation, is translated into a present position command, or PPC, that causes a synergy of arm muscles to contract and relax synchronously until the PPC equals the TPC. Figure 13.1 reviews the main variables that are computed by the VITE circuit.

Figure 13.1 shows that a TPC is subtracted from a PPC at a network stage at which a difference vector (DV) is computed. As discussed in Chapter 13, model DV properties are remarkably similar to data properties reported by Georgopoulos and his colleagues. In particular, a large DV may be active without causing an overt movement. Such DV activation, called *motor priming*, has been reported by Georgopoulos, Schwartz, and Kettner (1986). The next stage of the model (Figure 13.1) computes the product of a DV and a GO signal, or gain control signal, that controls read-out of the movement command. The product DV × GO is integrated through time by the subsequent PPC stage. Due to this circuit design, the PPC stage generates a continuous trajectory of synchronous outflow movement commands that gradually causes the PPC to approach the TPC at a speed that is regulated by the amplitude of the GO signal.

Since the completion of Chapter 13, a still larger body of behavioral and neural data about arm and speech articulator movements has been analysed by comparison with emergent properties of the VITE model (Bullock and Grossberg, 1988a; Cohen, Grossberg, and Stork, 1989). Here I summarize some neural data that support the existence of a GO signal, some predictions to further test the circuit's overall design, and a learning problem from which the circuit was derived to prepare the ground for comparative analysis with the vector map concept.

14.5. GO Signal Generator in Globus Pallidus

Additional physiological support for the VITE model comes from recent experiments involving lesions and electrical stimulation of the basal ganglia. Data from a set of experiments by Horak and Anderson (1984a, 1984b) are consistent with the interpretation that the internal segment of the globus pallidus is an *in vivo* analogue of the VITE model's GO-signal generator.

An *in vivo* candidate for a GO-signal generator must pass three tests. First, stimulation at some site in the proposed pathway must have an effect on the *rate* of muscle contractions. Second, it must have this effect without affecting the *amplitude* of the contractions. Thus stimulation should have no effect on movement accuracy. Third, this rate-modulating effect should be *nonspecific*: it should affect all muscles that are typically synergists for the movement in question.

The studies conducted by Horak and Anderson (1984a, 1984b) addressed these issues. Horak and Anderson (1984a) showed that "when neurons in the globus pallidus were destroyed by injections of kainic acid (KA) during task execution, contralateral arm movement times (MT) were increased significantly, with little or no change in reaction times (p.290)." This satisfies the rate criterion. Moreover, the rate of motor recruitment was depressed "in all the contralateral muscles studied at the wrist, elbow, shoulder, and back, but there were no changes in the sequential activation of the muscles (p.20)." This satisfies the non-specificity criterion. Finally, the authors also noted that "animals displayed no obvious difficulty in aiming accurately ... they did not miss the 1.5-cm target more often following KA injections, and there was no noticeable dysmetria around the target (p.300)." This satisfies the accuracy criterion.

Horak and Anderson (1984b) used an electrical stimulation paradigm instead of a lesion paradigm. They found that "stimulation in the ventrolateral internal segment of the globus pallidus (GP_i) or in the ansa lenticularis reduced movement time, whereas stimulation at many sites in the external pallidal segment (GP_e), dorsal (GP_i), and putamen increased movement times for the contralateral arm (p.305)." Once again, these effects were non-specific: "no somatotopic effects of stimulation were evident. If stimulation at a site produced slowing, it produced a depression of activity in all the muscles studied. Even stimulus currents as low as 25 μA affected proximal as well as distal muscles, flexor as well as extensor muscles, and early- as well as late-occurring activity (p.309)."

In the VITE model, activation of the GO-signal pathway produces movement only if instatement of a TPC different from the current PPC leads to the computation of a non-zero DV. In agreement with this property, Horak and Anderson (1984b) observed that "stimulation at sites that speeded movements did not induce involuntary muscle activation in resting animals nor did it change background EMG activity prior to self-generated activity during task performance (p.313)." In Chapter 13, it was noted that "very rapid freezing can be achieved by completely inhibiting the GO signal at any point in the trajectory". This property of the model has also

been shown to be a property of the GP system. In particular, Horak and Anderson reported that "stimulation with 50 or 100 μA at ... sites ventral and medial to typical GP_i neuronal activity completely and immediately halted the monkey's performance in the task (p.315)." Taken together, their experiments led Horak and Anderson (1984b) to conclude that "the basal ganglia ... determine the speed of the movement" (p.321).

14.6. Factorization of Position and Velocity Control

The striking correspondence between the experimental results of Georgopoulos *et al.* and of Horak and Anderson and the theoretical predictions of the VITE model regarding separate DV and GO-signal processes is important because it supports the hypothesis that motor systems, like sensory systems, implement a factorization of pattern and energy (Grossberg, 1970, 1978a, 1982a). In the motor system, this factorization means that a movement's velocity ("energy") can be scaled up or down over a wide range without disrupting the movement's direction or terminal position ("pattern"). Moreover, by using a GO-signal that grows gradually during the movement time, as in Figures 13.18 and 13.22, all synergists will complete their contractions at approximately the same time even if movement onset times of different synergists are staggered by a large amount (Bullock and Grossberg, 1988a, 1988b). These properties of the model, together with the growing evidence for separate DV and GO-signal pathways *in vivo*, provide a basis for understanding how primates can achieve space-time equifinality—all synergists reaching their length targets at equal times—yet retain separate control of velocity and position (Figure 13.24). Note that rate-control models relying on *static* stiffness adjustments (e.g., Cooke, 1980) lack the critical temporal-equifinality property.

14.7. Amplification of Peak Velocity and a GO Signal Test

Chapter 13 noted that in addition to compensating for muscles that begin to contract at staggered onset times, the VITE circuit automatically compensates for changes of target position during the movement time. In particular, it was shown that the model generates the amplification of peak velocity (Figure 13.23) that occurs during target-switching experiments (Georgopoulos, Kalaska, and Massey, 1981). Such a velocity amplification facilitates reaching the target after an incorrect initial TPC is replaced by an updated TPC. This speed-up occurs "on-the-fly." It is not preprogrammed, but is rather an automatic emergent property of VITE circuit interactions. It is caused as follows.

First there is a rapid change in the TPC and thus in the DV. Then a more gradual change occurs in the DV and the PPC as the PPC integrates the DV through time. The velocity amplification is predicted to be caused by interaction of the new DV with the GO signal that was activated by the previous movement command. This prediction can be physiologically tested using an experiment that would also provide another opportunity to test if the Horak-Anderson cells in globus pallidus generate GO signals. Such a test would directly stimulate, at increasing levels of intensity,

Horak-Anderson cells during a target switching task. Observed changes in movement velocity could then be used to calibrate the amount of change in GO signal amplitude caused by each level of stimulation.

14.8. Prediction and Test of Cells that Multiplex a Code for Local Velocity

The property of motor priming in motor cortex supports the VITE circuit prediction that there exists a stage subsequent to the DV stage at which an overt movement command is activated. The data of Horak and Anderson (1984a, 1984b) support the prediction that such activation may be partially controlled by a GO signal from the globus pallidus.

The VITE model may be further tested by recording from cells at the stage to which the motor cortex vector cells of Georgopoulos project. The cells at this subsequent stage should compute a measure of local movement velocity. This can be seen as follows.

These cells are predicted to compute (Figure 14.1) the product of rectified difference vector $[V]^+ = \max(V, 0)$ and GO signal G, namely

$$[V]^+ G. \tag{14.1}$$

The subsequent PPC stage computes a present position command P that performs a time integral of $[V]^+ G$, namely

$$P(t) = \int^t [V]^+ G dt. \tag{14.2}$$

Differentiating (14.2) shows that

$$\frac{dP}{dt} = [V]^+ G. \tag{14.3}$$

If subsequent network mechanisms cause the arm to follow closely the outflow movement command $P(t)$, then $\frac{dP}{dt}$ in (14.3) provides a good estimate of local movement velocity.

Equation (14.3) is of great interest from a conceptual point of view. It shows that the quantity $[V]^+ G$, which itself does not explicitly compute a velocity signal, *becomes* a velocity signal because of the manner in which it is transformed at subsequent network stages, in particular because it is integrated through time at the PPC stage.

The definition of this local velocity signal also depends upon the existence of a feedback loop between successive network stages (Figure 14.1). This feedback loop is needed to generate movements that are goal-oriented. This feedback loop is defined as follows. Quantity V in (14.3) computes a time-average of the difference between P and T, as in

$$\frac{d}{dt} V = \alpha(-V + T - P). \tag{14.4}$$

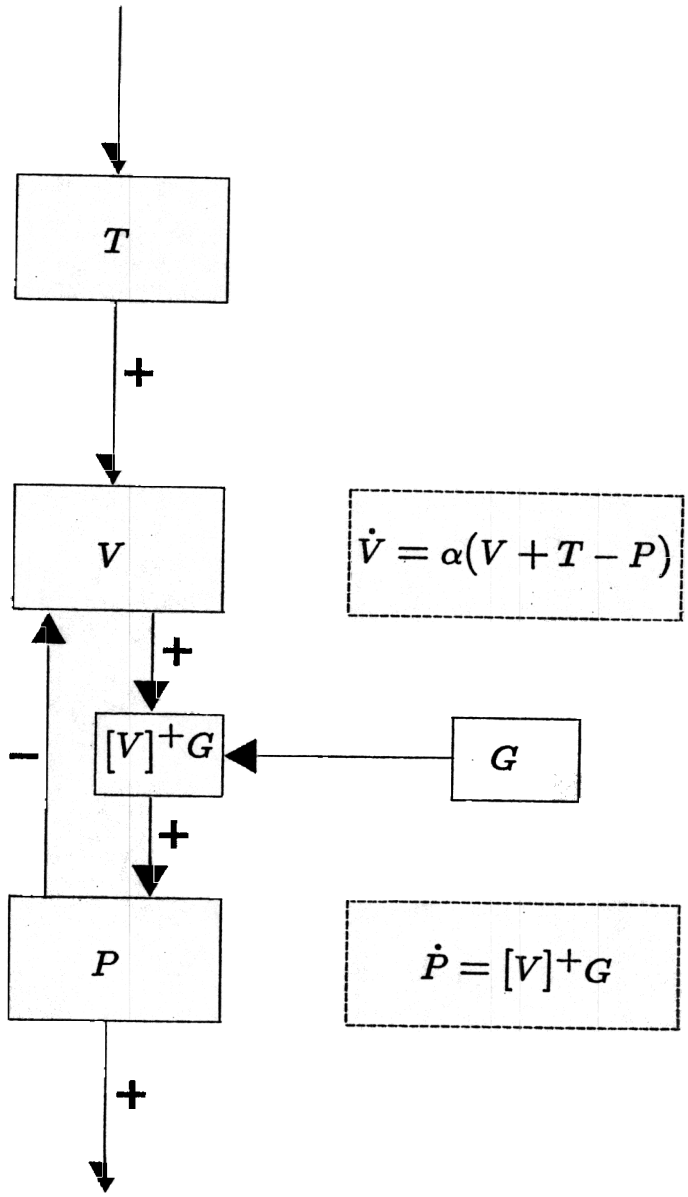

Figure 14.1. Main variables of the VITE circuit: T = target position command, V = difference vector, G = GO signal, P = present position command. The circuit does not include the opponent interactions that exist between the VG and P stages of agonist and antagonist muscle commands.

For simplicity, consider the case when the averaging rate α is large. Then, by (14.4),

$$V \cong T - P. \tag{14.5}$$

Hence, by (14.3) and (14.5),

$$\frac{d}{dt}P \cong [T - P]^+ G, \tag{14.6}$$

which summarizes the action of the feedback loop.

The approximate equation (14.6) shows that the local velocity signal is an emergent property of the entire network. It multiplexes quantities T, P, and G that are computed at three different network stages. It converts these quantities into a local velocity signal by using the feedback loop that exists between the DV stage and the PPC stage.

Given the motor cortical vector cells as an anatomical marker, it seems to be an experiment of great conceptual importance to test the existence of cells that code local velocity at one of their target nuclei. If the stage coding local velocity is found, then a retrograde marker applied at this stage may identify a location in, or near, the global pallidus. Then a direct neurophysiological test of the existence of a GO signal may be made to supplement the data of Horak and Anderson (1984a, 1984b). If globus pallidus cells are not marked, then this experiment may discover the true location of the GO signal generator. An anterograde marker applied at the local velocity stage may be used to discover the location of the PPC stage. Then a direct neurophysiological test can determine its ability to time-integrate local velocity signals.

14.9. Learning an Associate Map between Target Position Maps of the Eye-Head and Hand-Arm Movement Systems

Our theory suggests that the VITE circuit plays two distinct roles. One role, reviewed above, concerns trajectory formation within a sensory-motor system. This role may be called *intramodal trajectory formation.* The second role concerns *intermodal learning of target position maps* (Figure 13.14). The remarkable fact is that the same TPC, DV, and PPC stages are predicted to accomplish both roles.

The learning process transforms stored representations of a target position coded with respect to the eye-head system into a target position command of the hand-arm system for moving the arm, via VITE dynamics, to that position in space.

As pointed out in Section 13.14, not all positions that the eye-head system or the hand-arm system assume are the correct positions to associate through learning. For example, suppose that the hand briefly remains at a given position and that the eye moves to foveate the hand. An infinite number of positions are assumed by the eye as it moves to foveate the hand. Only the final, intended, or expected position of the eye-head system is a correct position to associate with the position of the hand-arm system.

Learning of an intermodal motor map must thus be prevented except when the eye-head system and the hand-arm system are near their intended positions. Otherwise, all possible positions of the two systems could be associated with each other, which would lead to behaviorally chaotic consequences. Several important conclusions follow from this observation:

(1) All such adaptive sensory-motor systems compute a representation of target position. This representation is the TPC.

(2) All such adaptive sensory-motor systems also compute a representation of present position. This representation is the PPC.

(3) During movement, target position is matched against present position. Intermodal map learning is prevented except when target position approximately matches present position, that is, except when the DV is small (Figure 13.14). A *learning gate*, or modulator, signal is thus controlled by the DV. This gating signal enables learning to occur when a good match occurs and prevents learning from occurring when a bad match occurs.

In summary, we trace the existence of vector difference codes to two fundamental computational problems: intermodal learning of target position transformations, and intramodal performance of synchronous trajectories within dynamically determined motor synergies.

14.10. Visually Reactive Movements and the Vector Map Code within Superior Colliculus

The type of vector map code described by Sparks and his colleagues may be analysed as part of a circuit design that realizes a different type of learning process. This learning process enables the *visually reactive movement system* to generate accurate reactive eye movements in response to flashing or moving lights on the retina (Chapters 2, 3, and 11). The role of map vectors in the visually reactive movement system has been analysed in several previous chapters as part of a developmental sequence during which reactive eye movements to flashing or moving lights on the retina are supplemented by attentionally mediated movements towards motivationally interesting sensory cues. These movements are supplemented once again by predictive eye movements that form part of planned sequences of complex movement synergies capable of ignoring the sensory substrate on which they are built. Each of these categories of eye movement requires one or more types of learning in order to achieve high accuracy. The movement systems wherein attention and intention play an increasingly important role base their adaptive success upon the prior learning of the more primitive, visually reactive types of movement.

14.11. Three Interacting Coordinate Systems: Retinotopic, Motor Sector, and Map Vector

The visually reactive movement system uses learning based upon visual error signals to improve the accuracy of its movements. It is assumed

that a target light is chosen and stored in short term memory (STM) before a movement starts (Figure 2.4). This stored representation activates a movement along an unconditioned movement pathway. The movement amplitude and direction, before learning, are controlled by a transformation of the retinal location of the stored representation into a motor representation in which the most eccentric positions generate the largest movement signals. This is accomplished by decomposing the motor representation into hemifields (Figure 2.1), and assuming that the gradient of connections to the corresponding pairs of agonist and antagonist muscles increases with eccentricity (Edwards, 1980; Gisbergen, Robinson, and Gielen, 1981).

The target light is stored in STM so that it can benefit from a visual error signal after the movement terminates. The stored representation samples this visual error signal along a conditioned movement pathway (Figure 2.4) whose output summates with that of the unconditioned pathway to generate the total movement signal. Thus learning within this system controls a feedforward adaptive gain that is changed by visual error signals.

How these visual error signals are coded clarifies one aspect of why a vector map exists in the deeper layers of superior colliculus. As illustrated by Figure 2.4, each light plays two roles: it acts as a movement signal for the next movement, and an error signal for the last movement. Thus the retinal location of each light must be remapped into a type of motor coordinates that can correct the full range of typical movement errors. The theory suggests that this is accomplished as follows.

14.12. Automatic Gain Control of Movement Commands by Visual Error Signals: Cerebellar Learning

The theory predicts how each movement command pathway can individually benefit from visual error signals to generate a more accurate movement in the future. Its analysis leads to a model of learning by the cerebellum which extends earlier models of cerebellar learning (Chapter 3). I emphasize two key properties of this model herein for purposes of comparative analysis: (i) the dual action of each light; and (ii) the learning of a motor synergy. The previous section summarized the network anatomy that subserves the dual action property.

14.13. Learning a Motor Synergy: Opponent Processing of Error Signals

The second key property of the AG stage concerns its ability to convert visual error signals, which individually activate only a *single* retinal position, into correct and synchronous movement commands to *all* muscles which move the eye. Learning of a motor synergy takes place in the Adaptive Gain Stage, or AG stage. The AG stage is identified with the cerebellar vermis, based upon data which show that this brain region controls modification of a saccade's *pulse gain* (Optican and Robinson, 1980). The conditioned movement pathway generates sampling signals which pass

through the AG stage and add or subtract a conditionable movement signal to the total movement command. An error signal acts to change the size, or gain, of the conditionable movement signal. Thus the AG stage is a region where automatic gain control of the total movement command takes place.

In order to learn a motor synergy, the system preprocesses the error signals before they can be sampled by the conditioned movement pathway. Two processing constraints conceptualize these preprocessing stages: (a) the Opponent Processing constraint, and (b) the Equal Access constraint. The need for Opponent Processing—which is a new feature of our model—can be seen as follows.

Each eye is moved by three pairs of agonist and antagonist muscles. One pair moves the eye horizontally. The other two pairs move the eye obliquely, and together can generate vertical movements (Figure 1.2). The model assumes that a learned *increase* in the gain of an agonist muscle command must generate a *decrease* in the gain of the corresponding antagonist muscle command, and conversely. In other words, each visual error signal has antagonistic, or opponent, effects on the conditionable gains of the muscle commands which it changes.

In order to realize this Opponent Processing constraint, suppose that the retina is topographically transformed from retinotopic coordinates into a motor map containing six sectors (Figure 2.2). This motor sector map is an idealization that may be compared to the data of Sparks and his colleagues. Each pair of agonist-antagonist muscles—(α^+, α^-), (β^+, β^-), (γ^+, γ^-)—is represented by opposite sectors in the sector map. A visual error signal which falls within a prescribed sector increases the conditioned gain of the corresponding muscle and decreases the conditioned gain of the antagonistic muscle. As shown in Figure 3.5 and the surrounding discussion, this type of retinal-to-motor transformation can be used to correct undershoot, overshoot, and skewed movement errors.

14.14. The Equal Access Constraint

In response to a light to a fixed retinal position, the system cannot *a priori* predict which type of error will occur as a result of its inadequately tuned parameters. In order to correct any possible error, each position must be able to activate a conditioned movement pathway that is capable of sampling error signals delivered to *any* of the motor sectors. This is the Equal Access constraint, which was first articulated in a formal model of cerebellar learning by Grossberg (1964, 1969a).

In order to realize the Equal Access constraint, we have assumed that the motor sectors are mapped, via a complex logarithmic map (Schwartz, 1980), into motor strips (Figure 3.9). Then a single conditioned movement pathway can sample gain changes due to error signals which activate any motor strip. Figure 3.6 describes two variants of this design. Each variant realizes both the Opponent Processing constraint and the Equal Access constraint.

The by now classical cerebellar interpretation of this anatomy (Figures 3.8 and 3.10) is that the sampling signals are carried by parallel fibers through the dendrites of Purkinje cells, whereas the error signals are carried by climbing fibers to the Purkinje cell dendrites (Albus, 1971; Grossberg, 1964, 1969a, 1972a; Ito, 1974; Marr, 1969).

In summary, visual error signals are mapped from retinotopic coordinates into motor sector coordinates and then into motor strip coordinates. Then any visual error signal can be sampled by any stored movement command within pathways organized to learn and read-out the correct conditioned gain signals to all the target muscles. Such a transformation provides a simple explanation of the type of vector map described by Sparks: An increase in map eccentricity increases movement length and a change in map polar angle changes movement direction because such a representation enables visual error signals to be mapped into motor commands that are capable of correcting undershoot, overshoot, and skewed errors in visually reactive movements.

14.15. The Vector-to-Sector Transform: Dimensional Consistency of Planned Vectors and Reactive Retinotopic Commands

It remains to discuss why the deeper layers of superior colliculus code movement *vectors*, rather than merely motorically transformed retinotopic commands. An analysis of this problem is spread over several chapters of this volume, namely Chapters 3, 4, 6, 10, 11. Here I outline some of the main design themes.

Perhaps the most salient issue concerns the apparent absurdity of using vector encoding when the problem is considered from a common-sense point of view. In order to compute a vector, the retinotopic location R of a light must be combined with the initial position E of the eye in the head to generate a target position T of the light in head coordinates. Let us symbolically represent this transformation by

$$T = R + E. \tag{14.7}$$

Then a vector V is computed by subtracting E from T:

$$V = T - E. \tag{14.8}$$

On the other hand, the eye moves in the head. Thus, all motor commands M must be recoded into head-coordinates again before activating the saccade generator that moves the eye:

$$M = V + E. \tag{14.9}$$

A comparison of equations (14.7)–(14.9) seems to suggest much ado about nothing, because (14.7) implies that

$$R = T - E, \tag{14.10}$$

so that M could have been derived directly from T without computing V at all!

The functional significance of these transformations is clarified by noting that vector coordinates V are consistent with retinotopic coordinates R, but head coordinates, T and M, are not. Visual error signals within the visually reactive movement system are retinotopically coded before being remapped into a retinally consistent motor sector map. In order for head-centered, attentive, planned movement commands T to benefit from the movement accuracy that is learned using visual error signals, they must be transformed into a retinally consistent coordinate system. Comparison of (14.8) and (14.9) illustrates that movement vectors V are consistent with retinal and motor sector coordinates.

Such vectors V are suggested in Sections 4.6–4.8 to be difference vectors, rather than the map vectors studied by Sparks and his colleagues. Thus both difference vectors and map vectors are suggested to exist as part of the eye movement control system. As with the difference vectors described by Georgopoulos and his colleagues for the control of arm movements, the difference vectors used in eye movement control are suggested to be computed in cerebral cortex (Sections 11.10 and 11.11).

In summary, the head commands T are transformed into vectors V so that the vectors V can be transformed into a motor sector code. In this way, attentive, planned movements can achieve the learned accuracy of visually reactive movements. This vector-to-sector transformation converts the motor sector map into a vector map.

14.16. Movement Gating, Intermodal Mapping, and Competition between Planned and Reactive Movements

The vector-to-sector transform is a learned transformation. Our model of how this learning process occurs (Sections 11.2 and 11.3) clarifies how vectors V can learn to control the movement pathways activated by a retinal position R during visually reactive movements when the vector V is generated by R, and thus $V = R$. The model also predicts how, after learning is over, a visually reactive movement to position R can be suppressed via a spatially organized competitive interaction when an intended movement command V is activated such that $V \neq R$. Despite this suppression, the model explains how the movement controlled by V achieves the accuracy derived from the cerebellar gains learned by visually reactive error signals. The same analysis also suggests how auditory signals (Section 11.4) and planned movement sequences (Chapter 9 and Sections 11.11 and 11.12) can activate accurate saccadic eye movements, and how inhibitory gating of superior colliculus by substantia nigra enables planned attentive movement commands to successfully compete with more rapidly processed reactive movement commands (see pages 131 and 286).

In summary, this comparative analysis clarifies the functional role played by two types of vector codes—vector map codes and vector difference codes—in the control of saccadic eye movements by the superior colliculus and the control of planned arm movements by the motor cortex,

respectively. Vector difference codes form part of the Vector Integration to Endpoint (VITE) Model that converts a target position command into a series of continuously integrated present position commands which are capable of generating a synchronous arm movement trajectory. The difference vectors are converted into overt movement commands by a gain control signal, called the GO signal, whose generator may be in globus pallidus. The VITE circuit is also predicted to play a role in modulating the learning of transformations from parietal target position representations of the eye-head system to target position commands of the hand-arm system.

The vector map codes are suggested to arise due to the interaction of several subsystems of the saccadic eye movement system. A visually reactive movement system uses visual error signals to correct motor synergies that are activated by visual signals. To accomplish this visual-to-motor learning process, the visual error signals are recoded into a motor sector map. An analysis of how movement errors are corrected by this system suggests a refined model of cerebellar learning, notably learning in the cerebellar vermis.

In order for the planned and attentive saccadic eye movement subsystems to benefit from visually reactive learning, they code their movement commands into vectors which are dimensionally compatible with the motor sector code, and then transform the vectors into the motor sector code. This vector-to-sector transformation converts the motor sector code into the type of vector map found in the deeper layers of superior colliculus. The properties of this transformation clarify how auditory signals and planned, attentive movement sequences may benefit from learning within the visually reactive movement system, and how gating of superior colliculus by substantia nigra enables the competition between these several movement systems to be successfully completed.

14.17. Data and Models of Posterior Parietal Target Positions Coded in Head-Centered Coordinates

The remainder of this chapter compares our theory with some of the relevant data and models that have appeared subsequent to the completion of the first edition of this book in 1985. When work on the book first began, few individuals were actively exploring self-organizing mechanisms of neural sensory-motor control. Now this literature is expanding rapidly. Thus the discussion herein does not aim for completeness. Rather its goal is to form some conceptual linkages that may be used by the reader to better understand the connections between our own results and those described by other investigators.

One area of great current interest concerns the existence and structure of a target position map coded in head-centered coordinates within the posterior parietal cortex. The need for such a map was impressed upon us by our discovery of system-level constraints concerning the manner in which visual targets coded in retinotopic coordinates could generate eye-movement commands coded in head coordinates. It hereby became clear

that combinations of visual signals and eye position signals would need to be combined in a distributed fashion in order to generate a head-centered target position map, or TPM. Chapter 10 described three related models in which invariant, but distributed, TPMs could be synthesized from combinations of non-invariant visual and eye position data. In the self-organizing model of this process (Sections 10.3–10.8), the initial position of the light on the retina and the initial position of the eye in the head are stored in short term memory before an eye movement begins. Early in the learning process, the eye movement is triggered by a visually reactive eye movement system. This system is sensitive to rapid changes in lights registered on the retina. After a movement is completed, a teaching vector $(I_1, I_2, \ldots, I_m)$ is derived from an outflow signal that characterizes the final position of the eye. However, the task of this TPM-learning system is to learn a correct *target* position, not just any final eye position of a movement.

Such accuracy is derived from the fact that the visually reactive movement system is sensitive to visual error signals that compute whether or not the eye movement enabled the eye to foveate the target light (Chapter 3). These error signals are used to trigger a learning process that enables visually reactive movements to become accurate. When these movements become accurate, the final eye position generates an internal representation of target position that is used as a teaching vector to learn an invariant TPM.

Our theory assumes that learning of an invariant TPM takes place while the system is making visually reactive movements. Thus an invariant TPM can be learned because the visually reactive movement system can correct its movement errors.

As reviewed in Section 14.15, target position commands within this TPM are transformed into vectors so that they can be associated, in a dimensionally consistent way, with retinotopically coded movement pathways within the visually reactive system.

14.18. Parallel Maps of Eye Position: An Application of Competitive Learning to Gaussian and Linear Teaching Vectors

In order to learn such an invariant TPM, two distinct representations of eye position information need to be computed in parallel and stored at different times during the eye movement cycle. The first eye position representation, which is denoted by EPM_I for *initial eye position map*, stores the position of the eye before the movement begins. This map is called EPM_1 in Chapter 10. A retinotopic map, which is denoted by RM, stores the retinotopic position of the target light before the movement starts. The second eye position representation, which is denoted by EPM_T for *target eye position map*, reads-out the eye position after the movement terminates. This map is called EPM_2 in Chapter 10. Both the RM and the EPM_I send sampling signals to the EPM_T. The EPM_T is the recipient of the teaching vector $(I_1, I_2, \ldots, I_m)$ after a movement terminates. Thus the model describes how an EPM_T can be adaptively transformed into

an invariant TPM, by building upon the learning of the (non-invariant) visually reactive movement system.

The theory assumes that both EPMs are derived from outflow signals that define the present position of the eye-in-the-head. These outflow signals are computed in motor coordinates that calibrate how much each of the six muscles that hold each eye in the head is contracted. Such a six-dimensional motor vector is transformed into an EPM via the mechanism of competitive learning (Sections 6.4–6.5). The most extreme form of competitive learning compresses a multidimensional vector into the choice of a single node, or cell population, in the EPM. Different nodes represent different motor vectors in such a map. A less extreme form of competitive learning converts a motor vector into a unimodal Gaussian activity profile within the TPM. Such a unimodal teaching vector for the EPM_T is used in one of the simulated models. See equation (10.49).

At the other extreme, the competitive learning mechanism does not distort the read-out of the motor vector. In effect, the motor vector is itself the teaching vector for the EPM_T. In this case, the teaching vector components I_k are (approximately) a linear function of eye position. This is a consequence of adaptive linearization (Chapter 5), which enables the muscles to respond linearly to motor commands, so that corollary discharges, in the form of EPM_T signals, can be used as an accurate teaching vector. This linear teacher defines the second model that was simulated. See equation (10.30). Also see equation (10.29), which assumes that the teacher is derived from an agonist-antagonist organization of eye position commands.

While the first edition of the book was being completed, Anderson, Essick, and Seigel (1984) presented some of their neurophysiological data about cells in posterior parietal cortex, and while the book was in press, Anderson, Essick, and Siegel (1985) published an article whose details supported the model's main predictions: that a distributed representation of target position in head-centered coordinates is implicitly generated by interacting combinations of visual and eye position signals, and that the receptive fields of the cells exhibit combinations of Gaussian and linear (or "planar") properties.

14.19. Back Propagation Model of Target Position: Comparison with Competitive Learning

Zipser and Anderson (1988) recently described a back propagation model of target position formation which built upon the data of Anderson, Essick, and Seigel (1985). This model assumes that RM, EPM_I, and EPM_T maps exist with Gaussian and linear receptive fields. It also assumes the existence of an extra intermediate level of "hidden units," or interneurons, between these maps in order to provide enough degrees of freedom for the steepest-descent curve-fitting mechanism of the algorithm to converge.

I do not believe that back propagation, at least in its present form, is a viable model of any brain process. This conclusion does not imply that

the algorithm may not be very useful in technological applications. The results of Zipser and Anderson (1988) illustrate my concerns about its use as a biological model.

First, there is a metatheoretical concern. With what probability will a model that was not derived to explain a particular set of physical or biological phenomena succeed in explaining it? For example, who would expect a theory of electrons to work if it was not derived from an analysis of data about electrons? At best, such a theory might capture the model-independent properties of electrons or, said in another way, the properties that have little to do with electrons.

Then there are a series of fundamental technical issues which too many proponents of back propagation, as a biological model, have heretofore either ignored or misrepresented. Sometimes this seems to have been the case because the practitioners do not understand how the model actually works. I have talked with scientists who have made use of convenient software tools to apply the model even though they do not know how it is defined.

This remarkable state of affairs has been justified by the claim that it does not matter how the model works. It is claimed that the model will generate the same "optimal" solution that a more biological model would also generate. Significantly, a similar type of claim was made about Artificial Intelligence models of brain processes when people worried that von Neumann computer architectures are different from biological brains. We now all realize that that claim was spurious.

The present claim as yet has no scientific justification. In fact, many properties of the algorithm strongly suggest that the claim is false and that back propagation will generate both quantitatively different results and qualitatively different heuristics than biologically motivated learning models.

For example, back propagation is, at bottom, a device for adaptive data compression. The number of hidden units is a key parameter in determining how compressed the data will become. This is, however, false in biological models such as competitive learning. There, as illustrated in Chapters 2, 6, and 10, an arbitrarily large number of hidden units may coexist with an arbitrary degree of data compression. The degree of compression is determined by how sharply the lateral inhibition, or competition, is tuned, not by how many hidden units exist.

This difference between the models leads to a second key difference. In back propagation, a single, teacher-determined, global error signal is broadcast throughout the network as a basis for changing its adaptive weights, or LTM traces. In competitive learning, there is no teacher, and all LTM weight changes are localized to the set of synapses that are contiguous to nodes that win the competition.

Perhaps the most unbiological feature of back-propagation is its use of a non-local transport of adaptive weights, or LTM traces, from its bottom-up filter to its top-down error signal (Figure 14.2). This computation lies at the heart of the algorithm. By this scheme, the numerical values of

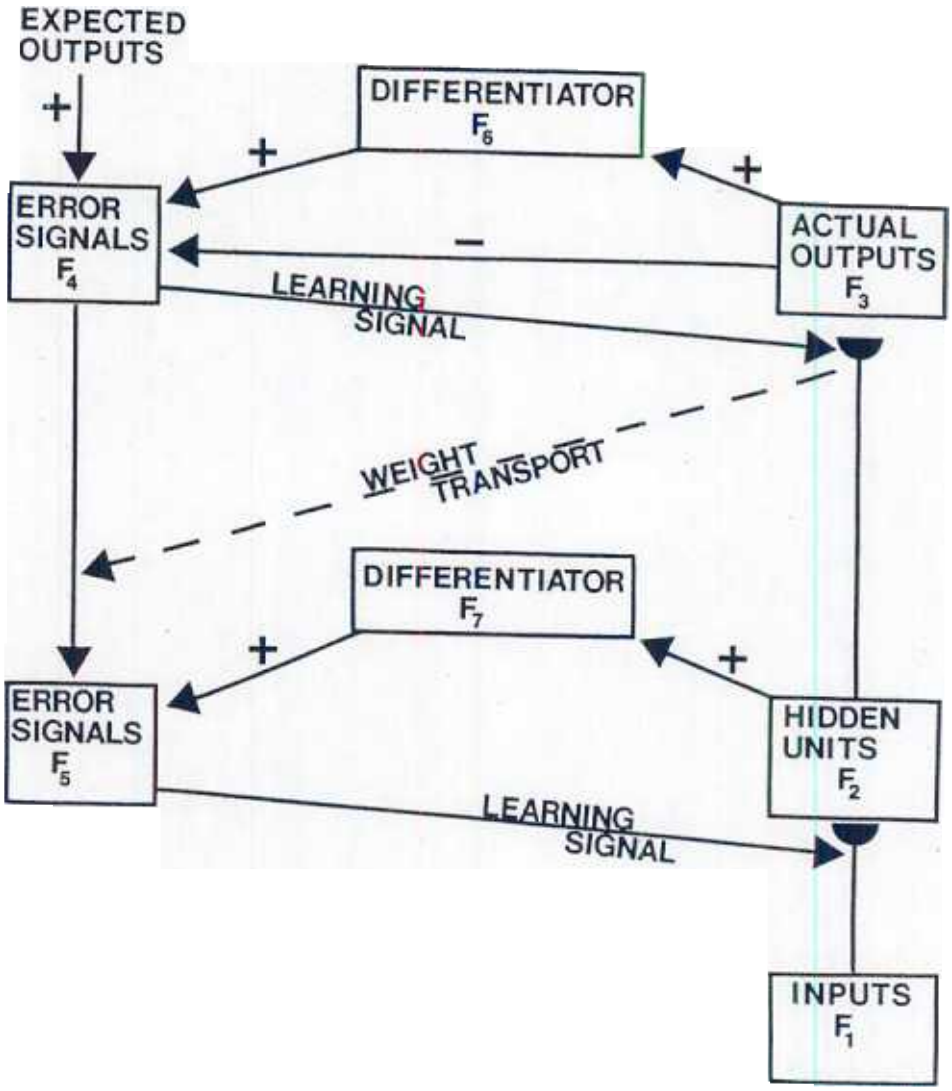

Figure 14.2. Circuit diagram of the back propagation model: In addition to processing levels F_1, F_2, F_3, there are also levels F_4, F_5, F_6, and F_7 to carry out the computations which control the learning process. The transport of learned weights from the $F_2 \rightarrow F_3$ pathways to the $F_4 \rightarrow F_5$ pathways shows that this algorithm cannot represent a learning process in the brain.

LTM traces that are computed within the $F_2 \rightarrow F_3$ pathway must be transported, with great precision, to distinct pathways $F_4 \rightarrow F_5$ elsewhere in the network. Such a non-local event has no support in neural data.

It is sometimes claimed that this weight transport operation somehow captures the familiar fact that reciprocal top-down pathways exist in all thalamo-cortical and cortico-cortical interactions. This claim is at best misleading. The top-down signals $F_4 \rightarrow F_5$ in Figure 14.2 do not directly influence the fast information processing at level F_2. Rather, they only have a slow and indirect effect on learning by the LTM traces in the $F_1 \rightarrow F_2$ pathways. Top-down, locally defined interactions which do influence fast information processing, as in Adaptive Resonance Theory, give rise to qualitatively different computational properties (Carpenter and Grossberg, 1987b, 1988b; Grossberg, 1987d). It is also well-known that the back-propagation algorithm may become seriously unstable when arbitrarily many inputs perturb it in real-time, when input statistics are nonstationary, or when learning rates are sped up to biologically plausible levels. Such problems do not occur, for example, in Adaptive Resonance Theory.

To sidestep these obstacles, some scientists point to the good fit of back propagation simulations to biological data. Unfortunately, no criteria of good fit are offered. One is invited to inspect a series of experimental receptive field properties and a series of simulated receptive fields which are acknowledged to contain significant discrepancies. In all examples that I have studied to date, even the qualitative fits are poor and typically worse than the fits achieved by alternative models in the literature.

In comparing the model of Zipser and Anderson (1988) with the models in Chapter 10, note in particular that the Zipser-Anderson model needs a level of hidden units, whereas the model of Chapter 10 does not; the Zipser-Anderson model uses non-local weight transport, whereas the model of Chapter 10 uses only local operations; the Zipser-Anderson model assumes Gaussian and linear receptive fields, whereas the model of Chapter 10 predicted such receptive fields. Further developments of such models should continue comparative analyses along these lines.

Perhaps the most unfortunate side effect of back propagation is that it seems to act as a mental soporific. At first, it seems to be so easy to use. As with all general curve-fitting techniques, one can always eventually manage to reduce the error by assuming enough degrees of freedom. Which degrees of freedom might be relevant to biological processing do not, however, seem to be suggested by the algorithm. As one who has always worked hard to find the kernel of truth in any work and then to run with it, I am still looking for signs that this technique can be used to generate new and nontrivial insights into neural representation.

14.20. Comparison of Mammalian and Anuran Head and Vector Representations

Cross-species comparisons and contrasts are of great importance for discovering computational invariants of the evolutionary process. A no-

table homology exists between the mammalian superior colliculus and the anuran tectum. Grobstein and his colleagues have made significant experimental advances and probing theoretical commentaries in their search for evolutionary invariants by using the frog tectum and related structures as an experimental model.

In their analysis of frog prey capture behavior, this group early noted the need for a target position representation coded in body-centered coordinates. In particular, Grobstein, Comer, and Kostyk (1981, p. 344) wrote that:

> "while a given retinal and tectal region corresponds to a single direction in an eye-centered co-ordinate frame, they in fact correspond to a set of directions in a body-centered or movement co-ordinate frame. Conversely, a given point in a movement co-ordinate frame in fact corresponds to a set of points in an eye-centered co-ordinate frame such as the retina or the tectum. We are intrigued by the possibility that between the tectum and the pattern generating circuitry there may be an intermediate level of circuitry into which space is represented in a body-centered or movement co-ordinate frame. Because of the many to one and one to many character of the transformation in going from an eye-centered to a body-center or movement co-ordinate frame, disturbances of circuitry involved in this transformation might be expected to result not in disconnection of particular tectal regions from pattern generating circuitry but rather in an alteration in the particular output associated with activation of given tectal regions."

More recent work has led Grobstein (1987) to posit the concept of an *activity gated divergence* as a path structure to link the retinotectal map with pattern generating circuitry. This concept embraces the several processing steps whereby a retinally coded visual input is transformed into a head-centered target position representation before being transformed once again into a vector map representation in the deep layers of the superior colliculus (see Sections 14.15–14.18). Using this interpretive bridge, one can make a link from the present book to Grobstein's data and his sophisticated commentary about other recent experimental and theoretical contributions.

14.21. The Transformation from Head-Centered Eye Movement Maps to Body-Centered Arm Movement Maps: Neck Corollary Discharges as a Map Teaching Signal

Chapter 10 described models of how retinally coded visual cues could be transformed into a target position map (TPM) in head-centered coordinates. Chapter 11 explained how this TPM could be used to generate saccadic eye movements towards a visual target, or towards a remembered target. Chapter 13 suggested how the same TPM could be used to generate visually-guided arm movements (Figure 13.14).

In order to carry out the transformation from an eye-head TPM to a hand-arm TPM, head-centered coordinates need to be transformed into body-centered coordinates by taking into account the position of the head in the body. Such a transformation can be learned by using convergent head-coordinate signals from the eye-head TPM and corollary discharges from the neck musculature to learn a representation of hand-arm target position with respect to the body. The model in Chapter 10 can be used to learn this transformation, just as it was used to learn the transformation from retinal to head-centered coordinates. Corollary discharges of arm position would replace the corollary discharges of eye position that were used as a teaching vector in Chapter 10.

In all, these considerations suggest that the same TPM that controls saccadic eye movements can be used to control arm movements in response to both momentary visual signals and remembered visual signals.

Nemire and Bridgeman (1987) have generated data which support this concept. Their studies confirm an earlier investigation by Gielen, van den Heuvel, and van Gisbergen (1984). In their experiments, forty large saccadic eye movements in darkness resulted in significant ocular undershoot of a remembered eccentric target position. Let us assume that this remembered target position was stored in short-term memory within the head-centered TPM. They also found a comparable amount of undershoot with manual pointing of the arm to the same remembered targets. Saccadic and arm movement measures were also highly correlated within sessions as well as between sessions. Because arm pointing could be changed by manipulating only saccades, the authors concluded that both systems share a single map of space.

14.22. Transformation from Auditory Maps to Visually-Activated Eye Movement Maps

The previous section discussed one type of intermodality transformation; namely, between the eye-head TPM and the hand-arm TPM. Both of these TPMs develop in response to visual cues. Sections 1.15 and 11.4–11.6 analysed how auditory cues can trigger saccadic eye movements. There it was noted that an incorrect eye movement in response to an auditory source does not generate an auditory error signal. In order to achieve accurate eye movements in response to auditory cues, it was predicted (pp. 22–23) that "Intermodality sharing of retinally activated saccadic command pathways can be achieved if there exists a processing stage at which signals generated by auditory cues feed into visually calibrated saccadic command pathways. Then auditory cues can use the visually learned saccadic parameters by activating these visually calibrated commands ... a learned mapping of auditory signals onto a visually derived head corodinate map would achieve the most parsimonious solution of this problem." This analysis thus led to the prediction that visual signals guide the learned adjustment of the map for auditory localization. See Figure 11.12.

Data from the laboratories of Knudsen and Konishi concerning the

existence of a head-centered auditory map in barn owls were cited (p. 22), but evidence was lacking concerning the main prediction that vision guides the adjustment of auditory localization.

While the first edition of this book was in press, experiments on barn owls that more directly supported this prediction were published (Knudsen, 1985; Knudsen and Knudsen, 1985). In these experiments, it was shown that barn owls raised with one ear plugged made systematic errors in auditory localization after the earplug was removed. These localization errors then corrected themselves within a few weeks in young owls, but not in adult owls, thus demonstrating a critical period of developmental plasticity. However, even young animals did not correct their auditory localization errors when deprived of vision. Moreover, when prisms were placed in front of their eyes, they adjusted their auditory localization to match the visual error induced by the prisms. Thus the visual system provides the spatial organization to which auditory localization is tuned.

14.23. A Related Model of Auditory-to-Visual Transformation: Neuronal Group Selection

Pearson, Sullivan, Gelfand, and Peterson (1987) responded to these Knudsen data by describing a model wherein auditory signals that are initially broadly distributed become more focused due to associative learning from the auditory map to the visual map. The model which they used is a variant of the competitive learning model that was used in this book for the same purpose.

However, instead of explicitly representing the competitive interaction in the network geometry, and the gating effects of the competition on associative learning, these authors used a lumped version of the competitive learning model. In this lumped model, the effects of the competition are directly built into thresholds for weakening or strengthening the auditory-to-visual LTM traces. This lumped competitive learning model is often called *neuronal group selection* in the writings of Edelman (1987).

14.24. Predictive Saccades and Saccade Sequences: The LTM Invariance Principle

Section 1.16, Chapter 9, and Sections 11.8–11.13 modelled a separate subsystem for storing predictive saccadic commands and sequences of commands in TPMs, for automatically reading-out these command sequences, and for adaptively calibrating the accuracy of predictive movements by using the visually reactive movement system as a source of accurate movement parameters.

Bronstein and Kennard (1987) have since reported additional data supporting the concept that the predictive saccadic movement and visually reactive movement systems are different.

Zingale and Kowler (1987) have reported parametric properties of predictive sequences of saccades. They noted that the latency of the first saccade in a sequence and the duration of intervals between subsequent saccades increased with sequence length, and varied with ordinal position

in the sequence. Our model (Section 9.4) for storing predictive saccadic commands in short term memory has all of these properties. It was derived to explain data about storing temporal order information in short term, or working, memory in a variety of perceptual, cognitive, and motor paradigms. The use of a similar circuit for all of these storage tasks was derived from the fact that all of these tasks face a common computational challenge: to store temporal order information in a way that permits its stable encoding into sequence codes, or chunks, in long-term memory. The design principle that constrains all such circuits is called the *LTM Invariance Principle* (Grossberg, 1978a, 1978c).

Zingale and Kowler (1987) noted the similarity of their data to data about typing, speech, and related motor tasks. The LTM Invariance Principle was used to explain such data in Grossberg (1986b).

14.25. Comparison of Saccade Generator Models

Sections 14.10–14.16 reviewed our theory of how a vector map code is generated in the deeper layers of the superior colliculus. The present section discusses recent modelling results relevant to our model of how vector commands from the superior colliculus input to the saccade generator (SG) to cause a saccadic eye movement. Section 2.5 and Chapter 7 develop this model.

Scudder (1988) has published a new model of the saccade generator. This model is an attempt to avoid the hypothesis of the Robinson (1975) model that a representation of target position is the input to the burst generator of the SG. Subsequent data, such as those of Mays and Sparks (1980) and Bruce and Goldberg (1985), have suggested that the superior colliculus and frontal eye fields generate inputs to the SG in the form of discrete vectors that do not change during the saccade. Scudder (1980, p. 1458) also makes this assumption; namely, that "the position of the eye relative to the target is evaluated once, the colliculus issues a command specifying the size and direction of the appropriate saccade, and the burst generator then executes it." Scudder contrasts his model with that of Keller (1979, 1980), in which the superior colliculus computes a motor error that continuously diminishes as the saccade progresses, a property that has not been supported by subsequent experiments.

In order to make a dimensionally consistent model, Scudder subtracts a vector specifying how much the eye has moved from the discrete vector that generates the SG movement command. This is accomplished by assuming that a type of inhibitory feedback neuron (IFN) exists in the SG that has not yet been experimentally discovered. The model also does not produce microsaccadic oscillations or saccade staircases.

Our own SG model shares a key property with the Scudder (1988) model; namely, we also assume that the superior colliculus generates inputs in the form of discrete vectors to the superior colliculus. On the other hand, our model does explain microsaccadic oscillations and saccade staircases, as well as such finer data features as antagonistic bursts near the end of a saccade, isometric coactivation during perpendicular saccades,

and saccadic undershoot during a fatigued state (Chapter 7). Our model achieves this greater explanatory range, moreover, without positing the existence of an IFN.

Instead, our model includes a postulate that partially supports the basic intuition behind the Robinson (1975) model, but also explains why a target position input to the SG has not been discovered. This is the postulate that an eye position representation, or EPM, adds to the vector representation, or RM, that is derived from the superior colliculus at the SG (Section 7.4). Both the RM and the EPM are switched on once and for all before the movement starts. The EPM is, however, normally cancelled by inhibitory feedback from the tonic cells (Figure 7.4). Thus its effects are not normally visible in the cell profiles of the SG. However, during the abnormal experimental conditions that generate saccade staircases, the EPM signal is not cancelled, and a saccade staircase is generated by the action of an Eye Position Update Network, or EPUN (Section 7.6). This prediction of our model may be used to differentiate it from the Scudder (1980) model by repeating the experiment of Schiller and Stryker (1972) to produce saccade staircases, while also recording from SG burst cells.

14.26. Applications of the Model's Circular Reaction to Eye-Hand Coordination by an Adaptive Robot

Kuperstein (1987, 1988a, 1988b) has used the models developed in this book to define a complete system for eye-hand coordination that he has implemented in a working robot. His system builds upon the concept of circular reaction (Section 1.3) and of the concepts developed in this book to computationally model such a sensory-motor cycle. Kuperstein also extends the models developed herein by defining representations of binocular eye position and binocular visual disparity and showing how to use these representations to generate accurate reaching movements to an object located in three-dimensional space.

Although Kuperstein has not yet incorporated all the models developed herein into his robot, his work provides an important existence proof that these models provide the foundation for "a novel approach to adaptive control of multijoint positioning which is at the boundary of both engineering and neuroscience. It is based on a new adaptive network control theory ... which allows a control system to learn a 'sense of space' by its own experience" (Kuperstein, 1987).