CHAPTER 1

MULTIPLE LEARNING PROBLEMS ARE SOLVED BY SENSORY-MOTOR SYSTEMS

1.1. Introduction: Brain Designs Are Adaptive Designs

One of the primary facts of life in the study of psychology and neurobiology is the remarkable multiplicity of behaviors, of levels of behavioral and neural organization, and of experimental paradigms and methods. One of the great needs in our science is to find unity behind this diversity.

This book describes a theory that unifies and predicts a large and diverse data base concerning the neural substrates of sensory-motor control. The book also illustrates a theoretical method that has unified other types of brain-related data using a small set of theoretical principles and mechanisms (Grossberg, 1982a, 1985a, 1985b).

The present work focuses upon the design principles and mechanisms whereby a particular type of sensory-motor system is controlled; namely, ballistic, or saccadic, eye movements. Although ballistic eye movements seem to be a relatively simple type of motor behavior, a large number of brain regions are utilized to control them, including retina, superior colliculus, parietal cortex, cerebellum, peripontine reticular formation, visual cortex, frontal cortex, and the oculomotor nuclei. The fact that such a simple type of behavior requires such a massive control structure has made the discovery of quantitative theories of brain dynamics difficult. This fact also raises serious issues concerning the specificity with which brain systems organize different types of behavior. If the brain does use specific types of circuitry, then why are so many different circuits needed to control even simple motor behaviors like ballistic eye movements? Moreover, how can a large number of circuits in a distributed control system coordinate specific and accurate behaviors?

The present approach focuses upon how brain systems are designed to form an adaptive relationship with their environment. Instead of focusing upon a few performance characteristics of a neural system, we consider the types of developmental and learning problems that a brain system *as a whole* must solve before accurate performance can be achieved. We have repeatedly found that an analysis of performance *per se* does not impose sufficiently many constraints to determine underlying control mechanisms. By contrast, an analysis of how development and learning lead to and maintain accurate performance characteristics has, time and again, opened a wide pathway to a rapidly expanding understanding of brain mechanisms. We believe that the unifying power of the theory is due to the fact that principles of adaptation—such as the laws regulating development and learning—are fundamental in determining the design of behavioral mechanisms.

Our analysis of ballistic eye movements has identified a set of distinct learning problems that its control system needs to solve in order to achieve accurate performance characteristics. The solutions of these learning problems take the form of real-time circuits that have a natural interpretation as neural networks. Even the simplest, or minimal, circuit solutions have been useful for organizing and predicting data concerning the different brain regions that control ballistic eye movements.

We have translated an anatomical multiplicity of brain regions into a functional multiplicity of learning problems. Behavioral, anatomical, and physiological data have been compared and contrasted with the minimal neural network circuits that are capable of solving these learning problems. We have crossed the conceptual gap between behavioral data and brain data by using these functionally meaningful networks as a bridge. A greatly expanded interdisciplinary data base could then be used to refine our understanding of the functional issues themselves.

With these networks in hand, one can better appreciate that the brain's solutions of its distributed control problems are both specific and efficient. Anatomical and physiological differences between brain regions can be analysed using network solutions of different developmental or learning problems. Using this approach, one can study how different sensory-motor systems solve similar learning problems by using the same functional characteristics that control ballistic eye movements. Different sensory-motor control systems can utilize specific circuits that pass through the same brain regions because these circuits all solve similar functional problems. Each brain region can thus be interpreted as a specialized functional processor that is shared by the many different circuits needing that functional capability. By tracing differences between sensory-motor skills to evolutionary variations on commonly shared functional designs, a significant compression of seemingly unrelated data can be achieved.

1.2. Eye Movements as a Model Sensory-Motor System

We have selected the mammalian eye movement control system to develop our theory because many workers are productively investigating this system as a model system for elaborating general principles and mechanisms of sensory-motor control (Baker and Berthoz, 1977; Fuchs and Becker, 1981; Ito, 1984; Zuber, 1981). Progress in understanding eye movements has been greatly accelerated by a fine tradition of quantitative modeling, inspired by workers like Masao Ito, David Robinson, and Larry Stark, within this field. Models of the eye movement system have rapidly progressed from formal control theory models towards neural network models (Robinson, 1973; van Gisbergen, Robinson, and Gielsen, 1981; Young and Stark, 1963) as the neural data base has expanded. Substantial conceptual progress has hereby been made towards understanding the performance characteristics of the neural components that control eye movements.

Recently it has also been increasingly appreciated that many eye movement characteristics can adaptively change. Robinson (1982) has, for example, reviewed adaptive properties of the control mechanisms of saccades, the vestibulo-ocular reflex (VOR), postural gaze, vergence, and balance behavior. A cerebellar circuit model has also been suggested to explain adaptive properties of the VOR (Fujita, 1982a, 1982b; Ito, 1982, 1984). However, adaptive models of other eye movements, notably saccades, have not been offered.

The present work grew in part out of a parallel development in the neural modelling literature. In Grossberg (1978a), some general principles and mechanisms of sensory-motor learning were articulated. The present work significantly extends this analysis within the specialized problem domain of the saccadic eye movement system. Due to our focus on learning issues, from the start we aimed at deriving a mechanistic understanding of how errors are corrected during saccadic learning. Such an approach rapidly leads to the realization that no individual neuron is able to assess the behavioral accuracy of an eye movement. The neural network as a whole needs to embody self-correcting mechanisms that can generate accurate behavior despite the ignorance of individual cells. Our central concern has been: How can neural networks learn accurate sensory-motor transformations even if the cellular parameters from which they are built may be different across individuals, may change during development, and may be altered by partial injuries throughout life?

One of the most difficult aspects of this work has been, and will continue to be, the identification of the conceptually distinct learning problems that a behaving organism must simultaneously solve in order for accurate reactive and planned movements to occur. The difficulty is due primarily to the one-to-many-to-one nature of behavior-brain relationships. Each individual sensory stimulus often generates an individual motor reaction. Between this ostensibly elementary one-to-one sensory-motor reaction lies a complex one-to-many analysis of the problem into several component parts before a many-to-one synthesis of the parts generates the observable motor reaction. Direct evidence concerning the nature of this one-to-many analysis and many-to-one synthesis can only be partially obtained using present experimental methods. Consequently, several of the basic problems which we have identified have not explicitly been described in the large literature of which we are aware. These component problems were not identified by piecing together the large mass of relevant behavioral and neural data fragments. They came into view through an analysis of the external visual and motor environment in which the eye movement neural system operates, and through actively confronting known eye movement data with known theoretical principles to test for matches and mismatches. As it turned out, in many of our theoretical circuits, known theoretical principles and mechanisms from Grossberg (1981, 1982a) could be developed, adapted, and refined to accomplish the requisite specialized tasks. This fact strengthens our conviction that there do exist general neural design principles and mechanisms, that some of them are already known, and that knowing them can enable a seemingly impossible set of problems to be reduced to a set of difficult but tractable problems.

As we have performed this reduction during the past six years, sensory-motor data and empirical models have been converging towards a greater appreciation of the importance of adaptive constraints upon neural designs. We therefore hope that this monograph will be viewed as a timely stimulus for focusing and amplifying efforts to understand the role which adaptive constraints play on the design of neural circuitry.

The previous discussion indicates why our theoretical analysis has been carried out simultaneously on several levels: behavioral and functional; anatomical; neurophysiological; and mathematical. The remaining sections of this chapter outline in intuitive language some of the major learning problems that need to be solved. Such intuitive language is not powerful enough either to identify or to solve all of the relevant problems. The language and concepts of nonlinear systems of differential equations are needed to do this. An intuitive description is, however, rich enough to clarify the nature of the problems and the scope of their solutions. With this intuitive description clearly in view, the reader can explore our detailed solutions of each separate problem, and its bearing on known and predicted interdisciplinary data, without losing the forest in the trees, or even worse, as we necessarily approach the level of a parametric mechanistic understanding, without losing the forest in the twigs and the leaves.

1.3. Intermodality Circular Reactions: Learning Gated by Comparison of Target Position with Present Position

Our starting point concerns an issue that is not discussed in the traditional eye movement literature (Grossberg, 1978a, Sections 48-51). This issue sets the stage for understanding the types of computations that are carried out within the eye movement system, as well as for understanding how eye movement commands can map onto movement commands for other sensory-motor modalities.

When an observer looks at an object, how does the observer's hand know where to move in order to touch the object? How is a transformation between the parameters of the eye-head system and the hand-arm system learned? Piaget (1963) has provided a deep insight into this learning process using his concept of a *circular reaction*. Imagine that an infant's hand makes a series of unconditional movements, which the infant's eyes unconditionally follow. As the hand occupies a variety of positions that the eye fixates upon, a transformation is learned from the parameters of the hand-arm system to the parameters of the eye-head system. To paraphrase Piaget's concept, we say that a circular reaction occurs when the *inverse transformation* from parameters of the eye-head system to parameters of the hand-arm system is also learned. This inverse map enables an observer to intentionally move its hand to a visually determined position.

How do the eye-head and hand-arm systems know what parameters are the correct ones to map upon one another? Not all positions which the eye-head system or the hand-arm system assume are the correct positions to associate. For example, suppose that the hand momentarily rests at a given position and that the eye quickly moves to foveate the hand. An infinite number of positions are assumed by the eyes as they move to foveate the hand. Only the final or intended position of the eye-head system is a correct position to associate with the position of the hand-arm system.

Learning of a circular reaction 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, thereby leading to a chaotic result.

How does the eye-head system know when it is close to its intended position? This discrimination can be made if the eye-head system can compute its present position, and can compare its intended position with its present position. The eye-head system has evidence that it is close to its intended position when a good match exists between these two types of information. Several important conclusions follow.

A. *Reciprocal Associative Transformations between Target Position Maps*

The first conclusion recapitulates and emphasizes the main insight of the circular reaction concept. Each adaptive sensory-motor system, such as the eye-head system and the hand-arm system, computes a representation, or map, of target positions. Such target positions may also be called intended positions or terminal motor positions, depending upon one's personal tastes.

B. *Matching of Target Position with Present Position*

Each adaptive sensory-motor system also computes a representation of present position. The target position is matched against the present position. During a movement, a fixed target position may be stored while changing present positions are matched against it. A central problem is to characterize the nature of this matching process. Target positions are *phasic* commands that can be switched on or off through time. For example, at times when a sensory-motor system is at rest, or in a postural mode, no target position whatsoever need be active in the system. By contrast, present position commands are *tonic* commands that are always on, since muscles are always in one or another position. How does one match phasic target position commands with tonic present position commands without causing spurious effects?

C. *Intermodality Map Learning is Gated by Intramodality Matching*

One such spurious effect is prevented by restricting the conditions under which intermodality learning takes place between target position maps (Figure 1.1). We assume that an active target position within the eye-head target position map can be associated with an active target position within the hand-arm system only at times when the target position approximately matches the present position. A *gating* signal is thus controlled by the network that matches target position with present position. This gating signal enables learning to occur when a good match occurs and prevents learning from ocurring when a bad match occurs.

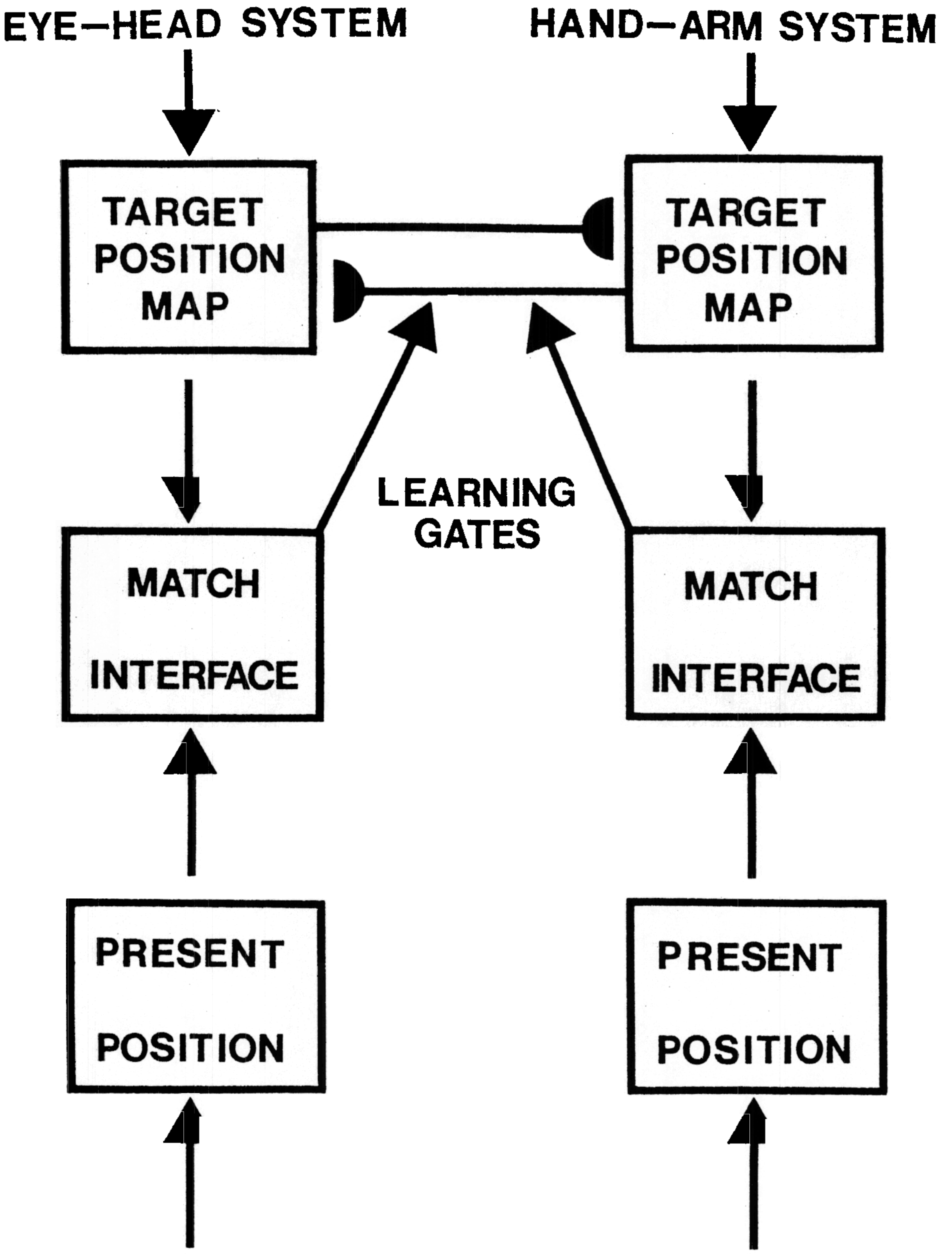

Figure 1.1. Learning intermodal circular reactions: Learning is gated by signals which are sensitive to how well target position matches present position within each modality.

D. *Dimensional Consistency: Head Coordinate Maps*

In order to compare target positions with present positions, both types of data must be computed in the same coordinate system. One cannot compare apples with oranges in neural networks any more than one can in any other scientific endeavor. Within a neural network, what one even means by a coordinate system is a deep issue that requires a systematic analysis. Some of the new problems that we have discovered have arisen from such an analysis. The conclusions of the previous paragraphs were derived without considering the coordinates in which the target positions and present positions are computed. An analysis of possible coordinates supplies important, and surprising, new information.

The present position of the eyes is computed with respect to head coordinates. By head coordinates we mean the following. The eyes rotate in their orbits with respect to the head. The directions in which the eyes point are determined by the amounts of contraction of the extraocular muscles (Figure 1.2). Signals either from the brain to the eye muscles (outflow) or from the eye muscles to the brain (inflow) could, in principle, be used to determine these directions. In either case, the eyes' present positions are computed relative to their position in the head.

As we noted above, in order to compare target positions with present positions, both types of data need to be computed in the same coordinates. Consequently, the target positions of the eye-head system are also computed in head coordinates. This conclusion has far-reaching implications.

1.4. Learning a Target Position Map

An analysis of intermodality circular reactions led to the conclusion that target positions of the eye movement system are computed in head coordinates. This section summarizes issues concerned with building up target position maps in head coordinates. When a light activates a retinal position, we say that it is registered in retinal coordinates. By Section 1.3, the position of a light in retinal coordinates is transformed into a target position computed in head coordinates in order to compute sensory-motor matches. Since this is true for all the possible retinal and target positions that can be activated by lights, we speak of transforming a retinal *map* into a target position *map*.

A. *A Many-To-One Transform*

A target position map is computed from combinations of visual and motor signals; namely, from the position of a light on the retina (visual) and the direction in which the eyeball is pointing in its head-anchored orbit (motor) before a saccade occurs. Many visual and motor positions correspond to a single target position. In other words, this transformation from retinal coordinates to head coordinates is many-to-one. For example, let a light hit the retina at θ degrees to the right of the fovea, and let the fovea point ϕ degrees to the right of its straight-ahead position (Figure 1.3). Many combinations of θ and ϕ correspond to the same target position

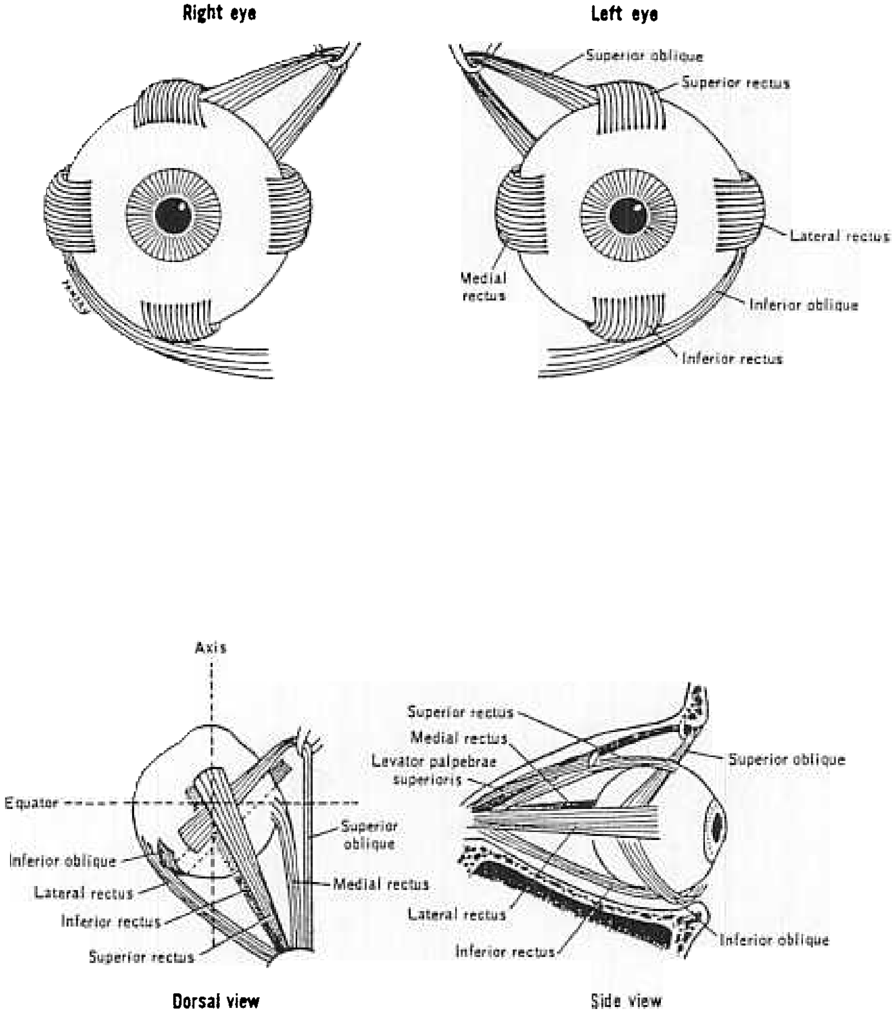

Figure 1.2. Each eye is moved by three agonist-antagonist pairs of extraocular muscles.

μ. A fixed value of μ determines a prescribed egocentric position of the light with respect to the head.

B. *Map Invariance*

This many-to-one relationship raises the general question of *map invariance*: How are command maps built up so that many combinations of input signals can all correspond to a single invariant map position? This question can also be asked in a language that is more familiar in linguistics or artificial intelligence: How can globally consistent *rules* emerge from locally ignorant components?

C. *A Multimodal Map*

Another aspect of the problem derives from the fact that we must transform "apples into oranges" to build up these invariants. That is, the signals to an invariant map often encode different kinds of input information (e.g., visual and motor) into yet another kind of output information (e.g., target position). The problem is to find a common dynamical language into which these diverse informational components can be expressed. Otherwise stated, the problem is to analyze invariant *multimodal* maps or rules.

D. *Error-Tolerance and Map Learning*

How can an *error-tolerant* invariant map get built up from such diverse types of information? The visual and motor systems of living creatures are constructed from many components, each of which may be error-prone. Individual differences in the parameters of these components can also occur due to fluctuations in developmental conditions or due to partial injuries throughout life. In order for an error-tolerant map to be generated from such variable components, some sort of self-organization, notably a self-correcting learning capability, is needed. Thus our problem is to design *self-organizing* invariant multimodal maps.

Invariance means that many combinations of visual and motor inputs can correspond to a single target position output. Expressed slightly differently, a *single* visual input is paired with many motor inputs to sweep out all the target position outputs to which that visual input contributes. A *single* motor input is paired with many visual inputs to sweep out all the target position outputs to which that motor input contributes. Within the context of a self-organizing system, this property of one-many pairing raises two more serious issues.

E. *Self-Consistent Map Learning*

The main issue concerns the possibility of learning a *self-contradictory rule*. A one-many pairing is needed to self-organize an invariant map, but how do each of these pairings encode its own correct invariant position without suffering from interference from all the other pairings in which each input participates?

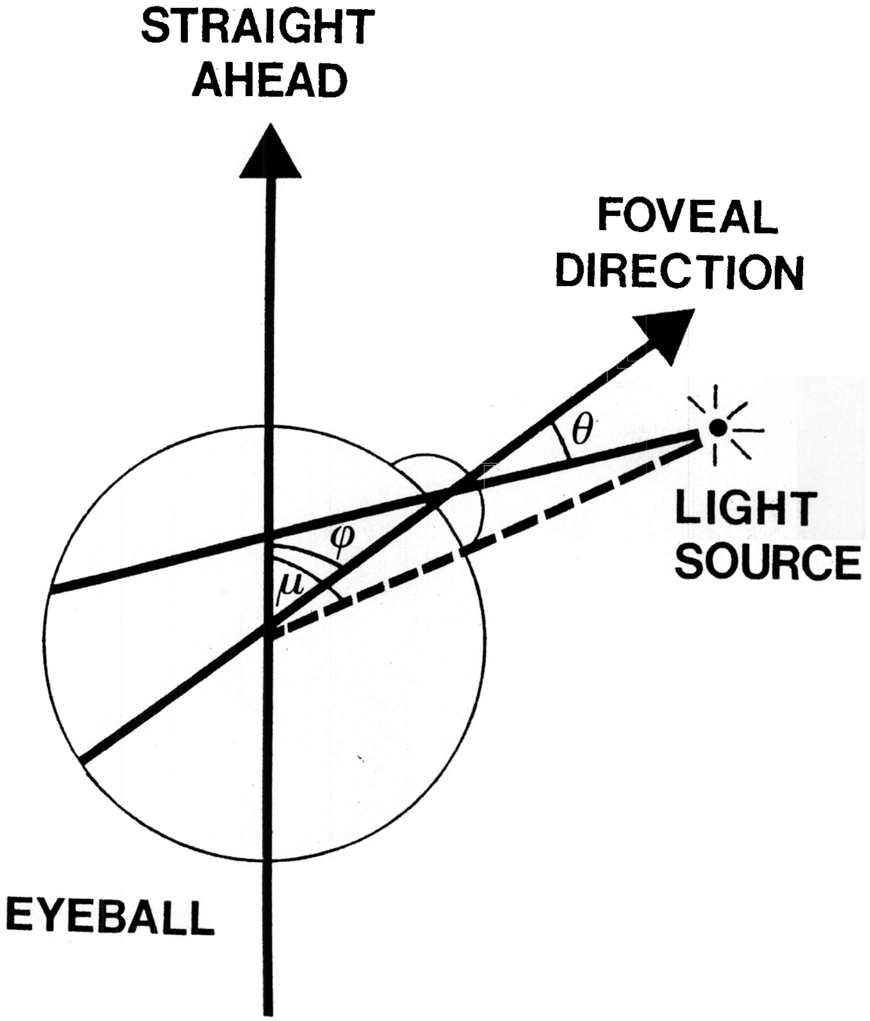

Figure 1.3. Geometry of eye position measured with respect to the head and retinotopic position measured with respect to the fovea: Angle ϕ measures eye position and angle θ measures retinotopic position. Angle μ measures position of the light source with respect to the head.

F. *A Self-Regulating Map*

Another aspect of this one-many issue concerns the question of "how many?". It would be absurd if a map could form in which each motor position is paired with 1000 visual positions, but as soon as it is paired with 1005 visual positions a self-contradiction arises within the system. The map should be able to form in response to a wide range of set sizes from which the inputs are drawn. Varying the set size spanned by the map inputs causes the set size spanned by the map outputs to also vary. Unless map self-organization is an unstable process, the invariant "form" of the map should be independent of its "size". This is the problem of *self-regulation*, which is one of the most important problems of developmental biology. Thus our problem is to understand self-organization of an invariant *self-regulating* multimodal map.

A finer aspect of the self-regulation issue concerns the number, or dimension, of input sources that combine to define a single intended position. We have spoken of a "visual input" and a "motor input". Each of these input channels can, however, be composed of several input pathways, or coordinates. For example, each visual input might be broken up into horizontal and vertical coordinates, or radial and angular coordinates, etc. Each motor input might be broken up into several individual motor inputs that correspond to all the components of the affected motor organ. Thus the self-regulation property includes not only the size of individual input fields, but also the number of input fields that cooperate to define each invariant position.

We will describe several different, but closely related, networks in which an invariant self-regulating multimodal map can form. Each of these networks is compatible with neural data and leads to testable predictions for future experiments. In such an invariant target position map, a single population v_μ can be activated by any of the many pairings of visual positions θ and of motor positions ϕ that correspond to the same target position μ. We will also show how a neural network as a whole can act *as if* it possesses an invariant target position map, whereas in actuality it only possesses maps of visual positions θ and of motor positions ϕ, or of noninvariant pairs (θ, ϕ) of all possible visual and motor positions. In order to understand how these different types of networks work, it is necessary to further analyze how target position commands are transformed into observable movements.

1.5. From Multimodal Target Map to Unimodal Motor Map

Information concerning one stage in this transformation can be acquired by considering the question of coordinates in greater detail. Section 1.3 suggested that target positions are compared with present positions in order to gate the learning of circular reactions. Target positions are partly derived from visual signals due to lights on the retina. By contrast, present positions of the eye within the head are described in motor coordinates. A fundamental calibration problem thus needs to be solved before the network can begin the comparison between target positions and present positions. A visually-derived target position needs to be transformed into motor coordinates so that it can be compared with present position signals that are also computed in motor coordinates. Obviously this transformation must be learned. Otherwise, we would have to conclude that all visual and motor parameters are calibrated with essentially perfect precision by a genetically pre-wired control of all developmental stages.

This learned transformation replaces one representation of target position (the visually-activated one) with another representation of target position (the motor one). Such a transformation involves both a change of map *coordinates* and a change of map *dimension*: The visually activated target position map possesses at least as many topographically distinct populations as there are discriminable lights on the retina. By contrast, the extra-ocular muscles, like many muscle systems, are organized into agonist-antagonist pairs. Each eye is moved by just three pairs of muscles. The present position of each eye can thus be characterized by six quantities. The transformation from visually-activated coordinates to motor coordinates replaces a large number of distinct map locations by arrays, or patterns, of six numbers. Such a transformation replaces complex and abstract combinations of multimodal information by simple and concrete arrays of unimodal motor information.

1.6. Vector Maps from Comparisons of Target Position Maps and Present Position Maps

The previous section noted that visually-derived target position commands need to be transformed into motor coordinates before they can be matched with present position commands that are also computed in motor coordinates. Our theory shows how the *same* network that learns this coordinate change *also* computes the match between target positions and present positions. We call this important network the *head-muscle interface*, or HMI.

The HMI has yet another important property. The degree of mismatch between a target position and a present position generates a motor code that can be used to accurately move the eyes. We conclude that a fundamental calibration problem that is solved through learning automatically generates properties that are necessary for skilled performance.

The degree of mismatch in the HMI represents a "vector difference" that compares the target position of the eye with its present position. This difference between target position and present position encodes how far and in what direction the eyes are to move. When the target position equals its present position, the vector difference equals zero, and no eye movement command is generated. By contrast, large mismatches between target position and present position represent commands to generate large saccadic movements.

1.7. Automatic Compensation for Present Position: Code Compression

The mismatches, or vectors, that are computed within the HMI harmonize two ostensibly conflicting design constraints. An analysis of circular reactions shows that *intermodality* sensory-motor commands are encoded as target positions. Target positions are not, however, sufficient to generate *intramodality* commands. In response to the same target position command, the eye needs to move different distances and directions depending upon its present position when the target position is registered. The HMI vectors automatically compensate for such changes in present position within each modality. They carry out this task in addition to gating associative learning of target position maps between modalities.

Two different types of data are manipulated by the HMI. Target position maps encode motor expectations or intended movements. They represent where the system wants to go. They are encoded and stored by the system long enough to execute the intended movement. They are thus switched on and off with an approximately digital, or logical, characteristic. By contrast, present position signals monitor the present state of the eye. Although they remain on tonically, they change continuously as the eye moves.

Using these two types of information, a *single* target position command can be rapidly transformed into *many* different movement trajectories due to the automatic compensation for present position that is encoded by the HMI vectors. This combination of mechanisms accomplishes a tremendous reduction in the number of commands that need to be stored. The network does not have to store many different movement trajectories with arbitrary initial and terminal positions, as inverse kinematic approaches to movement control would conclude (Brody and Paul, 1984). It does not have to map many different trajectories in one modality into many different trajectories in another modality. Instead, it computes maps of target positions. Such maps do not encode whole trajectories. They encode only the terminal positions of these trajectories. Terminal positions are mapped into terminal positions across modalities. Within each modality, terminal positions are mapped into motor coordinates via a learned transformation. By automatically compensating for present position, this transformation generates a code that can accurately control all possible movement trajectories.

1.8. Outflow versus Inflow in the Registration of Present Position

The need to solve additional learning problems becomes clear by considering possible sources of present position signals to the HMI. Two general types of present position signals have been acknowledged in discussions of motor control: *outflow* signals and *inflow* signals. Figure 1.4 schematizes the difference between these signal sources. An outflow signal carries a command from the brain to a muscle (Figure 1.4a). Signals that

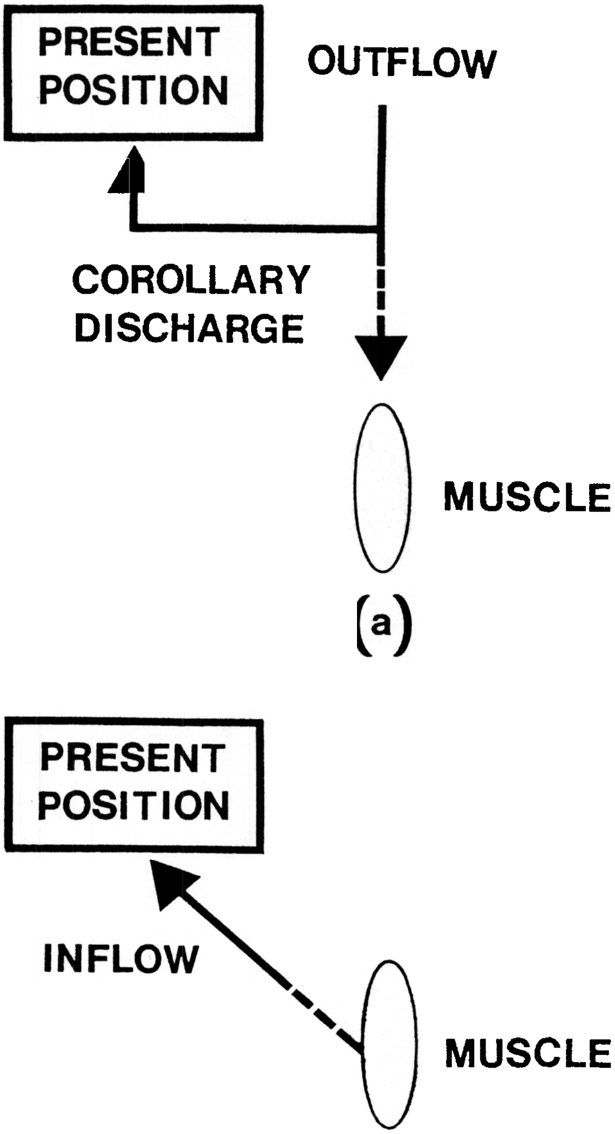

Figure 1.4. Outflow and inflow as sources of present position signals: (a) A source of outflow signals to a muscle branches to give rise to a corollary discharge; (b) A muscle gives rise to proprioceptive inflow signals.

branch off from the direct brain-to-muscle pathway in order to register present position are traditionally called *corollary discharges*. An inflow signal carries a command from a muscle to the brain (Figure 1.4b). Experimental evidence for both outflow and inflow involvement in motor control has accumulated over the years. Many of these tests aim to understand how the visual world achieves its apparent stability despite the fact that an observer's eyes are continually in motion (Epstein, 1977). A primary difference between outflow and inflow is that an outflow signal is triggered only when an observer's brain generates a movement command. By contrast, an inflow signal can be generated by a passive movement of the eye.

Helmholtz (1962) dramatized this difference by recommending that an observer jiggle his eyeball with a finger. The world seems to move due to these externally induced eye movements. By contrast, when an observer generates eye movements using brain-evoked signals, the world does not seem to move, even if an equivalent amount of visual movement is registered on the retina using the two procedures. This demonstration suggests that corollary discharges, rather than inflow signals, are used to compensate for self-induced movements in this situation. This conclusion follows from the fact that, if inflow signals are generated at all, they may be assumed to be generated whether an eye movement is internally or externally caused. Only when the eye movements are internally caused, however, does the brain compensate for the amount of visual movement that is due to the eye movement.

Although this type of demonstration strongly implicates corollary discharges as a source of present position signals, it does not imply that inflow signals play no role in the computation of present position. Since the pioneering works of Ruffini (1898) and Sherrington (1894), a large literature concerning inflow contributions to motor control has accumulated. Disentangling the different roles played by outflow and inflow signals has remained one of the major problems in the sensory-motor field.

1.9. Corollary Discharges and Calibration of Muscle Plant Contractions

Further insight about the roles of outflow and inflow can be achieved by considering the following facts. One role of an outflow signal is to move the eyes by contracting extraocular muscles (Figure 1.4a). However, the laws that govern the muscle plant are not known *a priori* to the outflow source. In particular, it is not known *a priori* how much the muscle will contract in response to an outflow signal of prescribed size. It is also not known how much the eyeball will rotate in response to a prescribed muscle contraction. Moreover, even if the system somehow knew this information at one time, it might turn out to be the wrong information at a later time. The muscle characteristics can change as they grow during development. They can also change as the body ages, or due to partial accidents, such as minor muscle tears or rupture of blood vessel capillaries.

These problems are serious ones even if the muscle contracts as a *linear* function of outflow signal size. A linear muscle plant's contraction is proportional to the size of the outflow signal. Linearity does not, however, guarantee that the proportionality constant, or *gain*, is neither tiny nor huge. What if the largest outflow signals could hardly move the eyes at all? What if even the smallest outflow signals could point the eyes in extreme nasal or temporal directions? A reasonable choice of muscle gain does not solve the problem of calibrating muscle responses if the muscle plant is nonlinear. Then the muscle contraction is not proportional to the outflow signal, so that doubling the outflow signal does not even double the amount of contraction, no matter how the gain is chosen.

These remarks illustrate that the relationship between the size of an outflow command and the amount of muscle contraction is, in principle, undeterminable without some type of additional information. This additional information must, explicitly or implicitly, characterize the muscle plant's response to outflow signals. To accomplish this, the system needs to compute a reliable measure of an outflow command as well as a reliable measure of the muscle's response to that command.

Corollary discharges provide a reliable measure of outflow commands (Figure 1.4a). In particular, corollary discharges are computed using signal sizes that the outflow pathway is capable of generating. The muscle responses to these signals may, however, initially be much too large, much too small, or even nonlinear, due to the characteristics of the muscle plant. In order to convert outflow signals into a full range of linear muscle contractions, somehow the brain needs to eventually adjust the responses of the muscle plant to these outflow signals. Such adjustments have the effect of causing the muscle to respond as if it were a different plant, notably a linear plant with a carefully chosen gain. From the start, outflow signals form a reliable basis on which to compute present position at the HMI. Inflow signals, whose plant characteristics are susceptible to continual change, do not.

1.10. Outflow-Inflow Pattern Matches and Linearization of Muscle Responses: Automatic Gain Control

The use of corollary discharges to compute present position at the HMI does not imply that inflow signals are not used. In fact, we argued in Section 1.9 that some type of information about muscle plant characteristics is needed to calibrate muscle contractions that veridically respond to outflow signals. We suggest that a brain region exists wherein comparisons between outflow and inflow signals are used for this purpose. We call this region the *outflow-inflow interface*, or OII.

The need for inflow data can be appreciated through the following arguments. How does the outflow system determine whether an outflow signal *should* cause a large or a small muscle contraction? Expressed in another way, how does the outflow system determine whether an outflow signal of a fixed size is large or small from a functional viewpoint? How big is "big"? What are the system's computational units?

An answer can be seen by recalling that outflow signals are computed in muscle coordinates, namely in agonist-antagonist coordinates. An outflow signal to its agonist muscle is "big" if the outflow signal to the corresponding antagonist muscle is "small." The *relative* sizes of agonist and antagonist outflow signals, not any absolute quantity, determine the desired size scale. Expressed in another way, the *spatial pattern*, or *normalized motor synergy*, of agonist and antagonist outflow signals determines the functional size scale.

If outflow signal sizes are computed in muscle coordinates, then the information which expresses the muscle plant's responses to these signals must also be computed in muscle coordinates. The simplest way to accomplish this is to use length-sensitive inflow signals from the muscles themselves.

This argument suggests that spatial patterns of outflow signals are matched against spatial patterns of inflow signals at the OII. Good matches imply that the muscles are responding linearly, and with a reasonable gain, to outflow signals. Bad matches must be able to adjust plant gain as well as plant nonlinearities. We will show in Chapter 5 how mismatches within the OII generate error signals that can change the size of the total outflow signal to the muscle plant. The conditionable part of the total outflow signal adds or subtracts the correct amount of signal to make the muscle react *as if* it is a linear muscle plant with a reasonable gain. The muscle plant does not itself change. Rather, automatic gain control signals compensate for its imperfections through learning. If the muscle plant changes due to aging or accidents, mismatches are caused within the OII and trigger new learning. The gain control signals automatically alter the total outflow command until the muscle again reacts linearly. Thus the linearization of the muscle plant is a learning process that takes place on a slower time scale than registration of a corollary discharge.

Throughout all of these learned changes, the corollary discharges to the OII remain intact. The system can compensate for plant changes without disrupting the code whereby present position is internally calibrated. Thus whereas outflow is used to rapidly change present position signals, inflow is used to drive slow recalibrations of the muscle response characteristics to these signals.

1.11. Motor Vectors Calibrated by Visual Error Signals

Linearizing the muscle plant's response to outflow signals does not ensure that the eye can move to accurately foveate a target light. Plant linearization just ensures that *if* outflow commands of correct size can be learned, then the muscles can faithfully execute these commands. Muscle inflow signals are certainly not the type of error signals that can determine whether the eye has foveated a target light.

A light on the retina is transformed by several processing stages before it can generate an outflow command to move the eye. Whether or not the eye successfully moves to foveate the light cannot be decided until the net effect of all these stages actually moves the eye. A reliable test of whether the light is foveated is given by the position of the light on the retina after the movement is over. Visual error signals are thus a reliable basis for modifying the sizes of outflow signals until the eye can successfully foveate target lights.

In Section 1.6, we concluded that a neural vector, by encoding the difference between a target position and the eye's present position, can be used to move the eye the correct distance and direction in order to foveate a retinal light. This conclusion did not specify how these vectors are transformed into correctly calibrated outflow commands. Two aspects of how this is done can now be noted.

The neural vectors within the HMI are encoded in muscle coordinates. Each neural vector is an activity pattern across a fixed set of muscle-coded cell populations. Changing neural vectors does not change *which* populations are activated. It only alters the patterning of this activity. How can a visual error signal *selectively* change the gains of different outflow commands if all of the vectors which generate these commands activate the same HMI populations?

A way is found by noticing that visual error signals are registered in retinotopic coordinates. Should not the HMI vectors be transformed into retinotopic coordinates in order to be dimensionally compatible with the visual error signals? If the HMI vectors are transformed from motor coordinates into retinotopic coordinates, then our problem is greatly simplified. Each motor vector would then excite a different location within a retinotopic map. Different retinotopically coded positions could control different outflow pathways, and their gains could therefore be separately altered by different visual error signals. A deeper reason for this vector transformation from motor coordinates to retinotopic coordinates will be described in Chapter 4.

When we say that a non-foveated light on the retina can act as a visual error signal, we mean that this visual signal somehow changes the outflow command into one that can generate a more accurate movement on the next performance trial. Such an outflow command controls a coordinated reaction, or synergy, of all the eye muscles. To solve this learning problem, one must therefore discover how a single light on the retina can alter an entire motor synergy. We will show in Chapter 3 how this can be accomplished. As mentioned above, a vector command is first transformed from motor coordinates into retinotopic coordinates. Then the vector command and its visual error signal are reconverted into motor coordinates in such a way that the visual error signal can correct the gains of all the relevant motor outflow signals.

1.12. Postural Stability: Separate Calibration of Muscle Length and Tension

The previous discussion suggests that several learning problems must be solved in order for the eyes to accurately foveate visual targets. Later chapters of the book will uncover and suggest solutions for subproblems and related problems that are not easily discerned without using a more mechanistic approach. All of these solutions are needed to understand how the eye moves reactively in response to retinal lights.

After the outflow command that determines a saccadic movement is over, how does the control system guarantee that the eye does not continue to move? What prevents the eye from drifting in its orbit in the absence of active saccadic commands? In other words, after movement is over, how is a stable posture assured?

This is a serious problem because the forces on the eyeball and on the extraocular muscles change rapidly as a saccadic movement terminates. Often the antagonist muscle relaxes while the agonist muscle is contracting (Bahill and Stark, 1979). This arrangement permits rapid motion to occur with a minimum of resistance from the antagonist muscle. By contrast, after the movement is over, the forces exerted by the agonist and antagonist muscles must be balanced in order to prevent large post-saccadic drifts from occurring. A new learning problem must thus be solved. It can be formulated as follows.

Any physically realizable position can be assumed by the eye muscles (Figure 1.5). Thus all possible combinations of agonist-antagonist muscle lengths can correspond to the final positions of saccades. At posture, the tensions of all agonist-antagonist muscle pairs must be equal. For *any* realizable combination of agonist-antagonist muscle *lengths*, how can *equal* tensions of agonist and antagonist muscles be generated in the postural mode?

This problem raises a number of new issues. How does the control system know when it is in the postural mode? How does it know whether the eye is drifting during this mode? If the control system could not tell the difference between the movement mode and the postural mode, then saccadic movements due to correct outflow commands could be misinterpreted as post-saccadic drifts, thereby leading to spurious "corrections" of already correct movement commands. What signals are used as the error signals to correct such a postural drift? How can these error signals balance the tensions of agonist and antagonist muscles without disrupting the correct lengths that were learned using visual error signals? We will conclude in Chapter 8 that visual error signals are also used to correct post-saccadic drifts. These visual error signals are not, however, error signals that compute how far a target light lies from the fovea. They compute the net motion of the eye with respect to the visual world during the postural mode. Visual information is, in fact, used in at least four different ways in our theory.

1.13. Planned versus Reactive Movements: The Rear View Mirror Problem

In Section 1.11, we concluded that visual error signals are used to calibrate the adaptive gains whereby the eyes can accurately foveate retinal lights. In order for this to happen, the saccadic control system must be sensitive to visual signals both before and after a visually reactive saccade occurs. We will now show why a different pattern of visual sensitivity

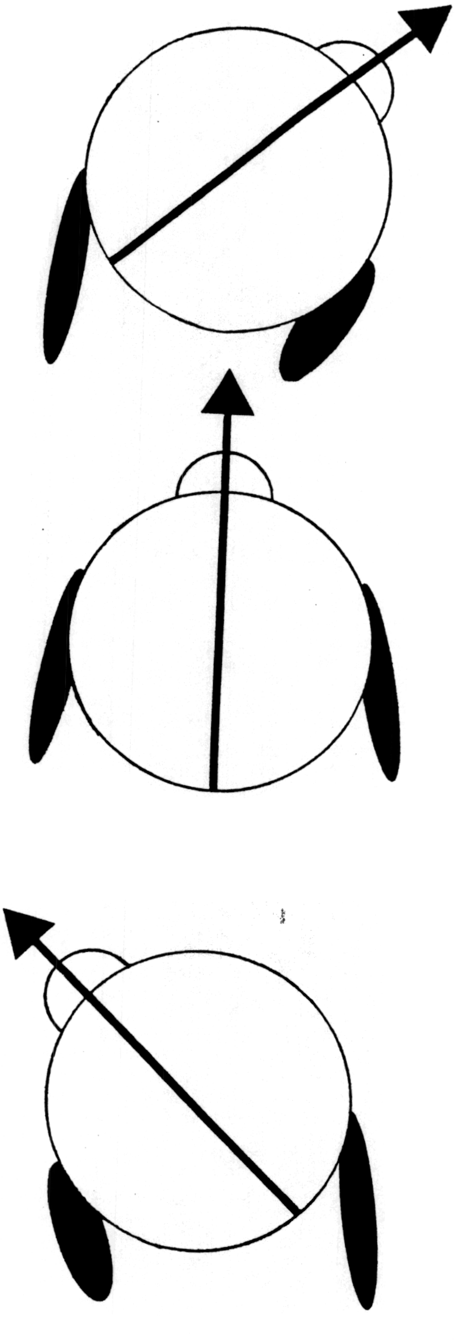

Figure 1.5. Different pairings of agonist-antagonist muscle lengths determine different positions of the eye with respect to the head.

exists when planned, or internally generated, saccades occur. Reconciling these two patterns of visual sensitivity requires additional learning circuits.

To understand this issue, imagine that you are driving your automobile, and that you wish to make a right turn when you reach a familiar corner. In order to determine whether it is safe to turn, you wish to look away from the roadway into your rear view mirror. Suppose that the dimensions of your body and car are such that looking into the rear view mirror can be accomplished by a saccadic eye movement.

Just before you make such a planned eye movement, your eyes are engaged by the flux of visual cues as seen through the windshield. In order to look away from the visual flux, your sensitivity to visual cues needs to be suppressed. Otherwise, the visual cues could command your attention. The initiation of a planned saccade thus requires the suppression of sensitivity to the visual cues that generate reactive saccades.

In order to make an accurate planned saccade, however, the adaptive gains that were learned in response to visual error signals must be used. We can now state the main issue: How do planned movements benefit from parameters that were learned during reactive movements, yet suppress the sensory signals that are needed to learn reactive movements? How can you have your parameters and suppress them too? This issue suggests that the command system which controls planned saccades is not the same as the command system which controls visually reactive saccades, yet that the reactive system can lend its learned parameters to the planned system *at the same time* that visual inputs to the reactive system are suppressed.

1.14. Attentional Gating

The distinction between planned movements and reactive movements raises the more general issue of attentional control. The decision to make a planned movement instead of a visually reactive movement involves a sensitivity shift that may be interpreted as a shift of attention. Even within the visually reactive system, many light sources can compete to be chosen as the targets for a saccadic movement. In order to understand how this occurs, we need to analyse how neural sensitivity can be modulated by sensory and learned factors, how a decision can quickly be made among many possible targets, and how all of these factors can be computed within the same coordinate system.

1.15. Intermodality Interactions in a Head Coordinate Frame

The last question brings us full circle back to the study of head, or egocentric, coordinates. Two types of arguments support the conclusion that attentional decisions take place within a head coordinate frame. The first argument is based upon the observation that sensory cues other than visual cues, notably auditory cues, can compete to be sources of saccadic movements. The second argument is based upon an analysis of the coordinate system within which planned movements are computed.

The first argument arises from a consideration of how intermodality sources of saccadic commands are adaptively calibrated. Section 1.11 pointed out that a retinally activated network for eliciting saccades is needed so that retinal fixation errors can be used to improve future saccadic accuracy via learning. How are accurate saccades in response to a source of sound, or to other nonvisual sources of sensory signals, generated? From what learning mechanisms do these alternative sources of saccadic commands derive their accuracy?

The contrast between visually and auditorily elicited saccades is particularly instructive. In the case of a visually evoked saccade, visual feedback can be used to correct saccadic parameters, because a target light on the retina can act as an error signal after the saccade occurs. This is true even if the head remains stationary throughout the saccade. A saccade to a sound source can also occur while the head remains stationary throughout the saccade. In this situation, however, no obvious source of auditory error signals exists, because the location of the auditory sound source does not change with respect to the head as a result of the saccadic eye movement. If no source of auditory error signals exists in this situation, then how do the eyes learn to accurately saccade towards a sound?

This lack of auditory error signals does not cause a problem if the auditory system can make use of the saccadic command pathways that have been adaptively tuned by visual error signals. Such an intermodality sharing of common pathways not only overcomes a problem of principle, but also significantly reduces the amount of adaptive machinery that is needed to improve the accuracy of intermodality saccades.

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. In order for auditory cues to effectively share visually tuned parameters, the overlapping intermodality maps must be dimensionally compatible. That is, if the maps represent coordinate systems with different invariant properties, then no consistent intermodality command structure can be learned. Thus in order to align intermodality coordinate systems, auditory and visual cues must be suitably preprocessed before they are mapped onto retinally activated pathways.

What type of auditory preprocessing is needed? Every auditory cue is registered with respect to head coordinates. An important example of such a head coordinate map has recently been worked out in the barn owl (Knudsen, 1984; Konishi, 1984; Sullivan and Konishi, 1984; Takahashi, Moiseff, and Konishi, 1984). By contrast, a light that is presented to a fixed retinal position determines a different position in head coordinates each time the initial eye position is changed. In order to align the two types of maps, either the auditory cue in head coordinates needs to be transformed into an auditory cue in retinotopic coordinates by compensating for changes in initial eye position, or the visual cue in retinotopic coordinates needs to be transformed into a visual cue in head coordinates. Visual cues are computed both in retinotopic coordinates and in head coordinates, in order to learn circular reactions (Section 1.3). Since auditory cues are directly registered in head coordinates, a learned mapping of auditory signals onto a visually derived head coordinate map would achieve the most parsimonious solution of this problem (Figure 1.6). Considerations other than parsimony will be used to strengthen this conclusion in the next section. However, the possibility cannot be ignored that different animals have evolved according to different evolutionary strategies in order to make intermodality map comparisons dimensionally compatible.

1.16. Head Coordinate Maps Encode Predictive Saccades

In order to ensure dimensional consistency, planned saccades and reactive saccades due to all relevant sensory modalities need to compete with each other within the same coordinate frame.

The need for a head coordinate map is also suggested by considering series of planned saccades, which we call predictive saccades, or the generation of predictive saccades within a more general motor program, such as a dance. A sequence of accurate saccades that occurs in response to the prior occurrence of a spatial or temporal pattern of lights illustrates predictive saccades that are responsive to controllable lights. For example, if a regular pattern of lights is briefly flashed in front of a human observer, the observer's saccadic eye movements can rapidly track the positions of the lights even after the lights are shut off (Kowler, personal communication). Hallett and Lightstone (1976) presented human subjects with two light flashes in sequence: one before a saccade and one during a saccade. In some cases subjects responded with two saccades after the second light presentation; the first to the first flash's spatial position and the second to the second flash's spatial position. Since the second saccade makes allowance for the size of the preceding saccade it cannot be based on retinal position alone. In what coordinate frame are the commands for such predictive saccades encoded and executed?

To clarify the main issues, suppose for the moment that each light in a pattern is encoded in a retinotopic map before the eye moves. After the eye moves in response to the first light, a correct eye movement to the second light can be made only if the eye movement system can compensate for the eye's prior saccade. This type of compensation is not just a matter of computing a new target position by taking into account the eye's new initial position before the second saccade occurs, because the retinal position of the second light is the wrong source to initiate a motion to the desired target position. Suppose, for example, that the eyes are pointing θ degrees to the right of the straight ahead direction when two lights excite the retina at ϕ_1 and ϕ_2 degrees to the right of the fovea (Figure 1.7). Suppose that the eyes somehow move to foveate the light at the ϕ_1 degree position, thereby attaining the eye position μ. No matter how the stored retinal position ϕ_2 and the approximate present eye position μ are juggled, one does not get the desired signal that can tell the

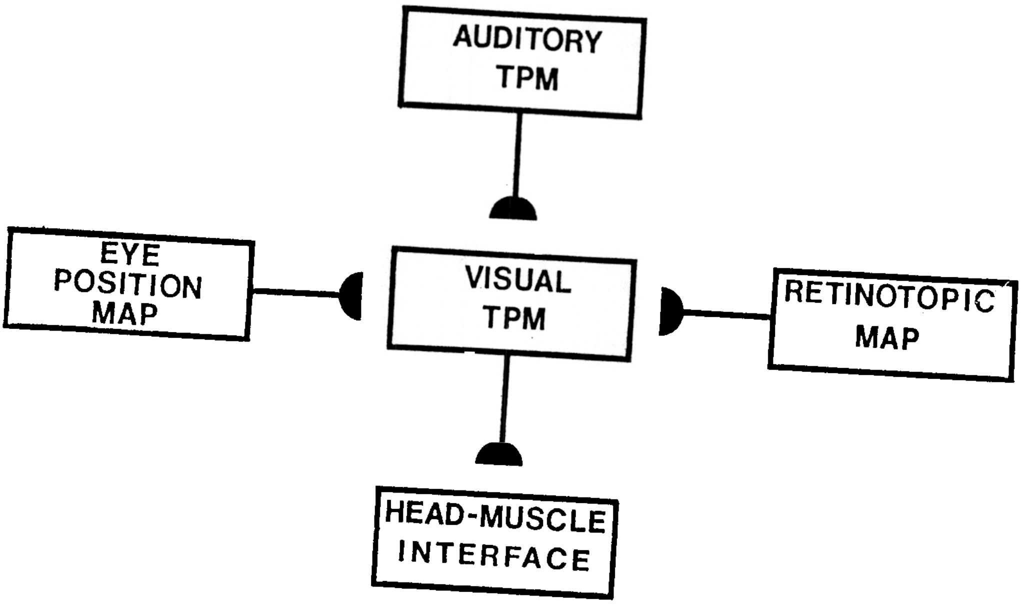

Figure 1.6. A visual target position map (TPM) receives inputs from several types of position maps before recoding its target positions into muscle coordinates at the head-muscle interface (HMI). A retinotopic map and an eye position map combine their signals at the visual TPM to implicitly define target positions there. Each TPM population can receive inputs from many retinotopic positions and eye positions. The auditory TPM can be consistently associated with the visual TPM because both are computed in head coordinates.

eyes how far they must move during the second saccade in order to foveate the second light. Both addition and subtraction of ϕ_2 and μ lead to the wrong answer.

One way to overcome this difficulty is to store the light positions in head coordinates before any saccade occurs. A predictive motion can be accurately made if the stored position of the second light with respect to the head is compared with the present position of the eye with respect to the head after the first saccade is over. The vector difference $\phi_2 - \phi_1$ of these positions determines the correct direction and length of motion that the eye must make to fixate the position of the second light. Then the process repeats itself. The position of the third light with respect to the head is compared with the position of the eye with respect to the head after the second saccade is over. And so on, until all stored saccadic commands are actualized.

In order for this argument to work, all the stored lights that will control a predictive saccade sequence must be stored once and for all in a head coordinate map. After the eye moves *any number of times*, the present eye position in the head can be subtracted from the target position, in head coordinates, that is coded by the next stored command. This "difference vector" represents the motion that will acquire the desired target location, if all possible calibration problems that are tacit in these statements can be solved.

These problems require the consideration of new issues. In addition to being able to simultaneously store all the predictive commands in a head coordinate map, the network that regulates predictive saccades must be able to store the commands in a way that reflects their temporal order, and must eliminate these command sources as their corresponding saccades are executed to make way for the next command in the series. A network that processes these storage, temporal order, reset, and calibration properties is described in Chapter 9.

Another possible way to encode difference vectors is to simultaneously store all retinotopic positions, such as ϕ_1 and ϕ_2, and to compute difference vectors $\phi_2 - \phi_1$ directly from the stored retinotopic values. Such a procedure seems very simple when it is stated without regard to how such vectors are recoded into movement commands expressed in motor coordinates. When such issues of coordinate transformation and calibration are studied, it emerges that a direct mapping from retinotopic values into difference vectors is more difficult to achieve than an indirect mapping from retinotopic values into head coordinates, followed by a comparison with present position. This problem is compounded by the fact that retinotopic coordinates are not invariant under eye movements, and are not suitable for intermodality comparisons, such as those which control eye-head and hand-arm coordination. We will therefore develop transformations from retinotopic into head coordinates, but will also describe in Section 9.10 the problems that need to be solved in order for direct recoding of retinotopic into vector coordinates to be effective.

The above discussion indicates that planned saccadic commands are

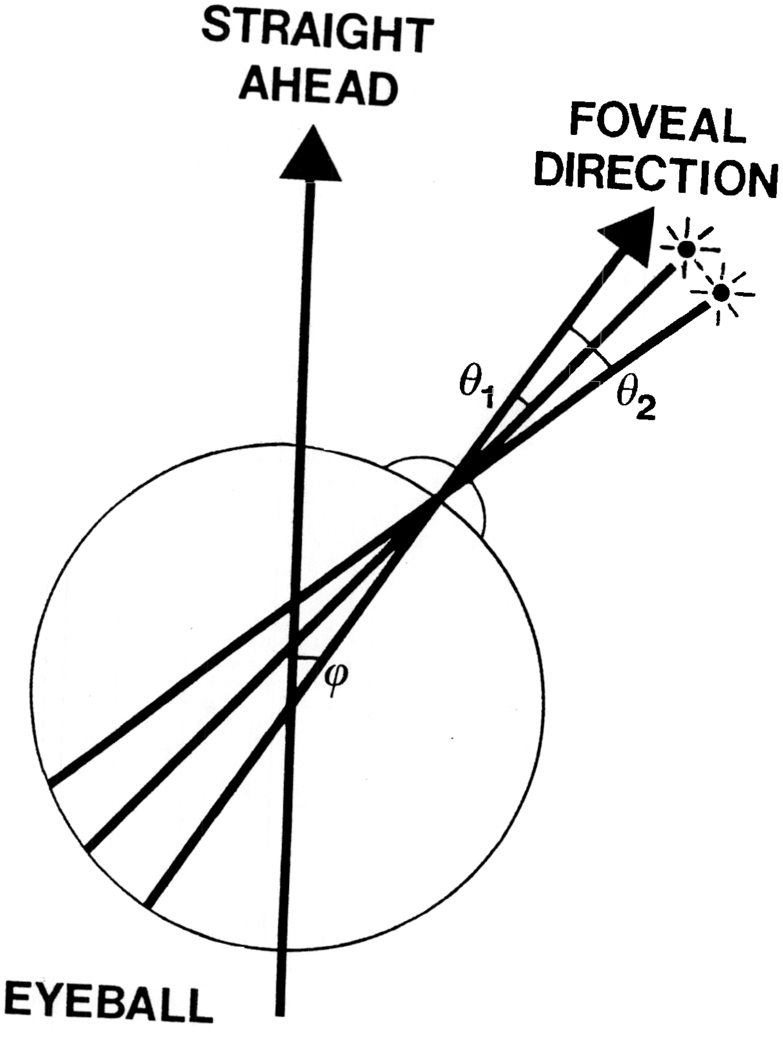

Figure 1.7. Two light sources, at retinotopic positions of θ_1° and θ_2° with respect to the fovea, are encoded in retinotopic coordinates while the eye points ϕ° with respect to the head.

initiated from command maps that are computed in head coordinates. Auditory sources of saccadic commands are directly entered into a head coordinate map due to the fixed positions of the ears in the head. When a saccade is initiated by sources of light, the transformation of retinal coordinates into head coordinates adds onto the retinal coordinates the effects of variable initial eye positions, and is thus analogous to a vector addition. The process whereby a target command in head coordinates is compared with present eye position transforms the head coordinate command into a retinotopic command. This transformation from head coordinates to retinotopic coordinates is analogous to a vector subtraction.

In all, a light that generates a planned saccade undergoes a coordinate transformation from a retinal coordinate map into a target position map in head coordinates before being reconverted into a retinotopic map via operations analogous to vector addition and vector subtraction, respectively. This reciprocal transformation, which seems quite pointless out of context, enables the saccadic system to make the decisions and calibrations that balance between visually reactive, intermodal, intentional, and predictive movement commands.

1.17. The Relationship between Macrotheory and Microtheory

The remainder of the monograph describes three types of information: the functional problems that are solved by an adaptive sensory-motor system such as the ballistic eye movement system; behavioral and neural data that are relevant to these functional problems; and the formal circuits that we propose as neural network solutions of the functional problems. In order to characterize the properties of these formal circuits, we use a combination of mathematical arguments and computer simulations.

The chapters are organized within the goal of maximizing the accessibility of the theoretical ideas. Each chapter therefore begins with a functional analysis and a description of relevant data. Technical details are developed towards the end of each chapter. Thus a feeling for most of the functional arguments can be acquired without the need to follow too many technical details. The mathematical definitions of the circuits are, however, needed to even state some of the concepts. This is true because of the following considerations.

Both the functional and the formal arguments need to be made on two levels: a Macrotheory level and a Microtheory level. These two aspects of the theory coexist in a mutually supportive relationship. The Macrotheory consists of several design principles, dynamical laws, and macrocircuits whose stages compute functionally characterized properties. The Microtheory describes the microcircuits that generate the properties of the various stages. Unlike many artificial intelligence models, the Macrotheory and the Microtheory cannot easily be dissociated. This is because the critical properties at the stages are interactive, or emergent, properties of the Microtheory's processes. Even the apparently local concept of feature detector is the net effect of widespread interactions within a Microtheory network.

The Microtheory thus does not freely invent properties at each macrostage. Each process of the Microtheory generates a formal grouping of interactive properties in response to prescribed external input and internal system constraints. The intrinsic structuring of these groupings defines the Macrotheory properties and is the source of the theory's predictive force. The Macrotheory's general principles and laws severely constrain the types of microprocesses that are allowed to take place at any stage. Only a few principles and laws are used in the entire theory, despite its broad intermodality scope.

A preliminary understanding of Macrotheory circuits can be arrived at from general functional arguments. Consequently, Macrotheory descriptions tend to accompany functional arguments at the beginning of each chapter. The actual transformations that solve the functional problems can, however, only be defined and understood by an analysis of Microtheory mechanisms. Thus an interplay of Macrotheory and Microtheory descriptions will necessarily permeate each chapter.

The macrocircuits that are suggested by this analysis lead to network subdivisions suggestive of such brain regions as retina, superior colliculus, peripontine reticular formation, cerebellum, parietal cortex, frontal eye fields, and oculomotor nuclei. The microcircuits that quantitatively instantiate the designs include analogs of such neuronal components as bursters, pausers, tonic cells, burst tonic cells, quasi-visual cells, Purkinje cells, mossy fibers, parallel fibers, climbing fibers, visuomovement cells, movement cells, light-sensitive cells, saccade cells, and post-saccadic cells. The properties of these circuits are used to suggest explanations of a large behavioral and neural data base in the subsequent chapters.

It remains to describe our strategy for discovering functional problems, and for ordering our exposition of these problems. We have used the *method of minimal anatomies* (Grossberg, 1974, p.69). "Given specific psychological postulates, we derive the *minimal* network...that realizes these postulates. Then we analyse the psychological and neural capabilities of this network. An important part of the analysis is to understand what the network cannot do. This knowledge often suggests what new psychological postulate is needed to derive the next, more complex network. In this way, a hierarchy of networks is derived, corresponding to ever more sophisticated postulates. This hierarchy...leads us closer to realistic anatomies, and provides us with a catalog of mechanisms to use in various situations. Moreover, once the mechanisms of a given minimal anatomy are understood, variations of this anatomy having particular advantages or disadvantages can be readily imagined." The method of minimal anatomies is thus a way to formally conceptualize successive stages in the evolution of a self-organizing system's ability to interact adaptively with different properties of its environment. By testing how simple networks succeed in adaptively solving some environmental problems but not others, a conceptual pressure is generated that points towards the designs which are needed to expand the network's total adaptive competence.

This chapter has provided a top-down view of many of the theoretical issues that the book analyses. In the subsequent chapters, we employ a bottom-up approach which is more suitable to the derivation of rigorous mechanisms. This bottom-up approach begins by considering networks that are much too simple to achieve a realistic competence. As we shall show, even these networks quickly point towards concepts capable of supporting a much more realistic theory.