Proceedings of the American Control Conference
Arlington, VA June 25-27, 2001

Self-Organizing Brain Dynamics by Which the Goals are Constructed that Control Patterns of Muscle Actions

Walter J Freeman

Department of Molecular and Cell Biology, LSA 129
University of California at Berkeley CA 94720-3200 USA
wfreeman@socrates.berkeley.edu http://sulcus.berkeley.edu

Abstract

Muscle control requires brain mechanisms that construct sequences of actions, predict each next step, and verify through reafference and proprioception that the desired movements are taking place. In a passive information processing system predefined stimuli convey information, which is transduced by receptors into trains of impulses that signify the features of an object. The symbols are processed according to rules for learning and association and are then bound into a representation, which is stored, retrieved and matched with new incoming representations. In active systems perception begins with the emergence of a goal that is implemented by the search for information. The only input accepted is that which is consistent with the goal and is anticipated in conjunction with the acts of searching. The key component to be modeled in brains is thedynamic subsytem that constructs goals and the adaptive actions by which the goals are achieved.

1. Neurodynamics of goal-seeking producing acts of observation

A device for measuring and recording such as a microphone, TV monitor or satellite can provide and store information, but it cannot make observations, nor can any extant computer to which the device is attached. The computer can be given examples of patterns that are to be sought, and it can be designed to detect the desired end states with specified levels of probability for positive and negative errors, but it cannot formulate the goals for itself, nor readily adapt the search strategy to its own changing needs and to rapid and unpredictable changes in its environment. Those are the characteristics of biological intelligence in animals, which are based in the capacities for self-organization in brain dynamics. Furthermore, voice- or motion-activated surveillance recorders can be triggered by prescribed inputs to limit their scope of intake, but once their gates are open, transcription is passive and representational, not active or meaningful. What is missing from these passive devices [Freeman 1999] is the goal-directedness contributed by observers.

The first step in understanding of the emergence of a goal in neurodynamics is to ask: what occurs in a brain during an act of observation? This is not a passive receipt of information from the world. It is a purposive action by which an observer directs the sense organs toward a selected aspect of the world, suppresses all of the input except that which is relevant, and interprets the resulting impact of the sensory barrage. The concept of intentionality has been used to describe this process in different contexts, since its first use by Aquinas (1272). The three salient characteristics of intentionality as it is treated here are (a) intent or directedness toward some future state or goal, (b) wholeness, and (c) unity [Freeman 2001]. These three aspects correspond to use of the term in psychology with the meaning of purpose, in medicine with the meaning of mode of healing and integration of the body, and in analytic philosophy with the meaning of the way in which beliefs and thoughts are connected with ("about") objects and events in the world.

Intent has 2 stages: first, the endogenous initiation, construction, and direction of behavior into the world, and second, the learning from the sensory consequences of the behavior. It is hypothesis testing, in which the outcome is predicted, and the observer learns from the test by action into the world. Humans and animals select their own goals, plan their own tactics, and choose when to begin, modify, and stop sequences of action, and humans at least are subjectively aware of themselves acting. Unity appears in the combining of input from all sensory modalities into *Gestalts (fused multisensory images)*, in the coordination of all parts of the body, both musculoskeletal and autonomic, into adaptive, flexible, yet focused movements. Subjectively, unity appears in the awareness of self. Wholeness is revealed by the orderly changes in the self and its behavior that constitute the development and maturation of the self, within the constraints of its genes and its material, social and cultural environments. Subjectively, wholeness is revealed in the remembrance of the self through a lifetime of change.

2. The limbic system is the organ of intentional behavior

Brain scientists have known for over a century that the necessary and sufficient part of the vertebrate brain to sustain minimal intentional behavior is the ventral forebrain, including those components that comprise the external shell of the phylogenetically oldest part of the forebrain, the paleocortex, and the deeper lying nuclei with which the cortex is connected. These components suffice to support remarkably adept patterns of intentional behavior, in dogs after all the newer parts of the forebrain have been surgically removed, and in rats with neocortex chemically inactivated by spreading depression. Intentional behavior is severely altered or absent after major damage to the medial temporal lobe of the basal forebrain, as manifested most clearly in Alzheimer's disease. This same part of the brain is essential for the formation of declarative, episodic memories that are subject to conscious recall [Clark and Squire, 1998], as distinct from procedural memories such as those formed by classical conditioning procedures.

Phylogenetic evidence comes from observing intentional behavior in salamanders, which have the simplest of the existing vertebrate forebrains [Herrick 1948; Freeman, 1995]. The three parts are sensory (which, as in small mammals, is predominantly olfactory), motor, and associational. The latter part contains the primordial hippocampus, which is identified in higher vertebrates as the locus of the functions of spatial orientation by the "cognitive map" [O'Keefe & Nadel, 1978] and temporal integration of learning (the organization of long and short term memory [Figure 1]).

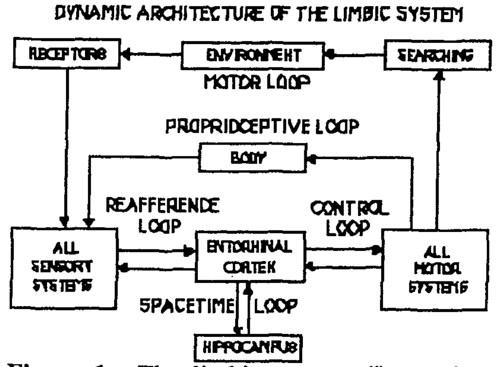

Figure 1. The limbic system ("spacetime loop") gets all sensory input and distributes it to all sensory and motor systems

For example, hunger is an emergent pattern of neuroactivity that expresses the requirements of brains and bodies for metabolic fuel and building material. It induces a state transition in the neural populations of the forebrain under the influence of sensory stimuli from the gut and the brain's own chemoreceptors for its chemical state. It is also shaped by neurohormones from nuclei in the brain stem. The emergent pattern impacts the brain stem and spinal cord, leading to stereotypic searching movements that are adapted to the immediately surrounding world. Feedback from the muscles and joints to the somatosensory cortex provides confirmation that the intended actions are taking place. The impact of the movements of the body on sensory input is conveyed to the visual, auditory and olfactory systems.

All of the perceptual constructs, which are triggered by sensory stimuli and dependent on prior learning, are transmitted to the limbic system, specifically to the entorhinal cortex, where they are combined. When an animal detects an odor of food, it must hold it, move, take another sniff, and compare the two concentrations in order to decide which way to move next. The difference in strength has no meaning, unless the animal has records of which way it moved, when the samples were taken, and a basis for determining distance and direction in its environment. These basic operations of intentional behavior are properties of the limbic system. The same requirements hold for all distance receptors, so it is clear why evolution has provided for prior multimodal sensory convergence in order to perform space-time integration on the Gestalt, not on its components.

3. Properties of EEG activity patterns

These brain states are observed in spatial patterns of EEG (electroencephalographic) potentials that are not representations of stimuli, nor are they the simple effects caused by stimuli. Each learned stimulus serves to elicit the construction of a pattern that is shaped by the synaptic modifications among cortical neurons from prior learning, and also by the brain stem nuclei that bathe the forebrain in neuromodulatory chemicals [Freeman 1992, 2000]. It is a dynamic action pattern that creates and carries the meanings of stimuli for the animal. It reflects the individual history, present context, and expectancy, corresponding to the unity and the wholeness of the intentionality. The patterns created in each cortex are unique to each animal. All sensory cortices transmit their signals into the limbic system, where they are integrated with each other over time, and the resultant integrated meaning is transmitted back to the cortices in the processes of selective attending, expectancy, and the prediction of future inputs by reafference.

The same kinds of EEG activity as those found in the sensory and motor cortices are found in various parts of the limbic system. This discovery indicates that the limbic system also has the capacity to create its own spatiotemporal patterns of neural activity. They are related to past experience and convergent multisensory input, but they are self-organized. The limbic system serves to generate the neural activity that forms goals and directs behavior toward them. EEG evidence shows that the process occurs in rapid steps, like frames in a b/w motion picture. Each frame is formed by a 1st order state transition from a basal receiving state.

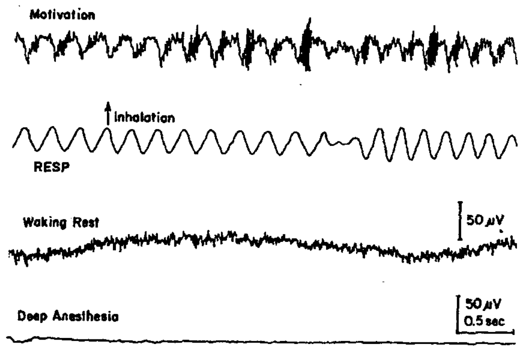

Figure 2. EEGs in deep anesthesia (open loop, bottom, 4th trace), resting (3rd trace), and aroused (top). Bursts of oscillation result from destabilization of the system with each inhalation (upward in 2nd trace).

Each step follows a dynamic state transition, in which an assembly of neuron populations jumps suddenly from one spatiotemporal pattern to the next, as the behavior evolves. Being intrinsically unstable, the limbic system continually transits across states and broadcasts them to other parts of the brain, and then relapses to a basal state so as to give rise to new ones. Limbic output controls the brain stem nuclei that serve to regulate its excitability levels, implying that it regulates its own neurohumoral context, enabling it to respond with equal facility to changes that call for arousal and adaptation or rest and recreation, both in the body and the environment. The neurodynamics of the limbic system, assisted by other parts of the forebrain such as the frontal lobes, initiates the novel and creative actions seen in search by trial and error.

4. Characerization of brain states

In neurodynamics [Freeman 2000] analysis begins with defining three kinds of stable state, each with its type of attractor. The point attractor holds the neurons at rest unless perturbed by input, and they return to rest when allowed to do so. Examples of point attractors are silent neurons or populations that have been isolated from the brain or depressed by a strong anesthetic, to the point where the EEG has gone flat (Figure 2). This is the 'open loop' state in which the time and space constants of populations are measured from the linear dendritic response to impulse input, $\delta(t)$, with rates of rise, a, and decay, b, from fitting the curve with the solution to a 2nd order ODE, where k_{ij} is coupling for input to dendritic synapses, v is wave, p is pulse density:

$$d^2v/dt^2 + (a+b)dv/dt + ab\ v = k_{ij}G(p) \quad (1)$$

and G(p) is an asymmetric sigmoid function governing wave-pulse conversion. This static nonlinearity is derived from the Hodkin-Huxley equations, for the asymmetric wave-to-pulse conversion applied to populations:

$$p=p_o\,(1+\{1-\exp[-(e^v-1)/Q_m]\}),\quad v > -u_o \quad (2)$$
$$dp/dv = u_o \exp\,[v-(e^v-1)/Q_m] \quad (3)$$

Instability is made input-dependent by the asymmetry. It creates an attractor landscape composed of the basins determined by learning classes of input.

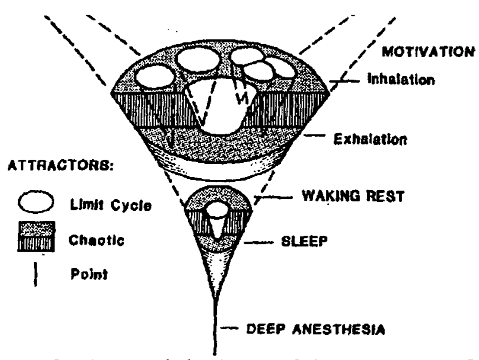

Figure 3. A portrait is shown of the state space of the olfactory system. In a motivated state the system oscillates between chaotic and limit cycle attractors recreating and dissolving the landscape rapidly to allow repeated sampling of odors.

A special case of a point attractor is noise. This state is observed in populations of neurons in the brain of a subject at rest, with no evidence of overt behavior (Figure 2, 3rd trace). The neurons fire continually but not in concert with each other. Their pulses occur in long trains at irregular times. Knowledge about the prior pulse trains from each neuron and those of its neighbors up to the present fails to support the prediction of when the next pulse will occur. The state of noise has continual activity with no history of how it started, and it gives only the expectation that its amplitude and other statistical properties will persist unchanged.

A system that gives periodic behavior is said to have a limit cycle attractor. The classic example is the clock. When it is viewed in terms of its ceaseless motion, it is regarded as unstable until it winds down, runs out of power, and goes to a point attractor. If it resumes its regular beat after it is re-set or otherwise perturbed, it is stable as long as its power lasts. Its history is limited to one cycle, after which there is no retention of its transient approach in its basin to its attractor. Neurons and populations rarely fire periodically, and when they appear to do so, close inspection shows that the activities are in fact irregular and unpredictable in detail, and when periodic activity does occur, it is either intentional, as in rhythmic drumming, or pathological, as in nystagmus and Parkinsonian tremor.

The third type of attractor gives aperiodic oscillation of the kind that is observed in recordings of EEGs. There is no one or small number of frequencies at which the system oscillates. The system behavior is therefore unpredictable, because performance can only be projected far into the future for periodic behavior. This type was first called "strange"; it is now widely known as "chaotic". The existence of this type of oscillation was known to Poincaré a century ago, but systematic study was possible only recently after the full development of digital computers. The best known simple systems with chaotic attractors have a small number of components and a few degrees of freedom, as for example, the double-hinged pendulum, and the dripping faucet. Large and complex systems such as neurons and neural populations are thought to be capable of chaotic behavior, but proof is not yet possible at the present level of developments in mathematics.

The discovery of chaos has profound implications for the study of brain function (Haken 1983; Skarda and Freeman 1987). A chaotic system has the capacity to create novel and unexpected patterns of activity. It can jump instantly from one mode of behavior to another, which manifests the facts that it has a landscape of attractors, each with its basin, and that it can move from one basin to another in an itinerant trajectory (Tsuda 1996). It retains in its pathway across its basins its history, which fades into its past, just as its predictability into its future decreases with distance. Transitions between chaotic states constitute the dynamics we need to understand how brains do what they do.

5. The cortical state transition is the elemental step of intentionality

Systems such as neurons and brains that have multiple chaotic attractors also have point and limit attractors. A system that is in the basin of one of its chaotic attractors is legendary for the sensitivity to what are called the "initial conditions". This refers to the way in which a simple system is placed into the basin of one of its attractors. If the basin is that of a point or a limit cycle attractor, the system proceeds predictably to the same end state. If the basin leads to a chaotic attractor, the system goes into ceaseless fluctuation, as long as its energy lasts. If the starting point is identical on repeated trials, which can only be assured by simulation of the dynamics on a digital computer, the same aperiodic behavior appears. This is why chaos is sometimes called "deterministic". If the starting point is changed by an arbitrarily small amount, although the system is still in the same basin, the trajectory is not identical. If the difference in starting conditions is too small to be originally detected, it can be inferred from the unfolding behavior of the system, as the difference in trajectories becomes apparent. This observation shows that a chaotic system can create information in the course of continually constructing its own trajectory into the future. In each sensory cortex there are multiple basins corresponding to previously learned classes of stimuli. and also the unstimulated state.

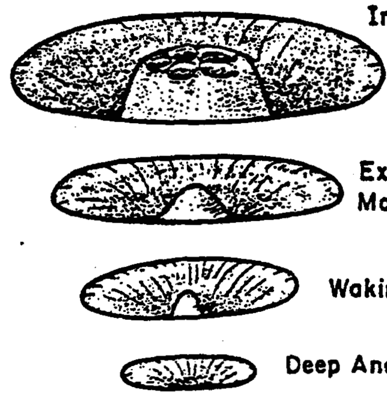

Figure 4. A perspective projection of the state space from Figure 3 shows the learned attractor landscape that forms with each inhalation, giving the basis for rapid selection by a stimulus, and that then dissolves, clearing the entire landscape to allow receiving and classifying the next sample from the environment.

Together they form an attractor landscape that repeatedly forms and dissolves at each new stimulus frame. The chaotic prestimulus state in "exhalation" (Fig. 4) establishes the sensitivity of the cortex, so that the very small number of sensory action potentials evoked by an expected stimulus can carry the cortical trajectory into the basin of an appropriate attractor. The part of the world in which the stimulus is to be sought is selected by the limbic brain through orientation of the sensory receptors by sniffing, looking, and listening. The basins of attraction are shaped by limbic input to sensitize the reception of input belonging to a desired class of stimuli. The web of synaptic connections that was modified by learning forms the landscape of basins and attractors that serve to guide acts of observation.

Access to attractor basins is by destabilization of the cortical neural populations, manifested in a 1st order state transition. The attractor by which the system is then governed determines the spatial pattern of the oscillatory output by imposing amplitude modulation on the shared wave form, which is transmitted by axons in parallel using pulse frequency modulation and time multiplexing to an array of targets, where the signal is extracted after smoothing by a spatial integral transform imposed by a divergent-convergent axon projection.

The selection of the basin of attraction is done by the stimulus given to the receptors feeding the primary receiving cortex. The basins and their attractors are shaped by Hebbian and non-Hebbian learning, as well as by reafferent barrages [Kay et al. 1996; Kay and Freeman 1998] that bias and tilt the attractor landscape. The trajectory manifested in the EEG is higher in dimension than the landscape, so that every basin is immediately accessible from every point of the trajectory with minimal need for search. Capture by one of the attractors reduces the dimension of activity below that of the landscape.

6. Conclusions

We are clearly still a long way from full understanding of intentionality in animals, and even farther from constructing analog systems that will embody the principles of active perception by which brains achieve their ascendency. The main inference that I wish to propose from these biological studies is that there is nothing mysterious or inaccessible about intentional behavior, and that given the right insights into how simpler brains organize themselves in the necessary construction and pursuit of their own goals, there is no insuperable barrier to building devices that have those capacities. There is already substantial progress in that direction through the development of embodied robots [Clark, 1996; Hendriks-Jansen, 1996; Brooks 1999]. It is not too soon to begin to consider the ethical issues that success in these projects will surely raise.

Acknowledgments

This work was supported by grants from the National Institute of Mental Health, and from the Office of Naval Research.

References

Aquinas, St. Thomas (1272) The Summa Theologica. Volume 19 in the Great Books Series. Chicago:Encyclopedia Britannica. Inc., 1952.

Barrie JM, Freeman WJ, Lenhart M (1996) Modulation by discriminative training of spatial patterns of gamma EEG amplitude and phase in neocortex of rabbits. Journal of Neurophysiology 76: 520-539.

Brooks, R.A. (1999) Cambrian Intelligence: The Early History of the New AI, Cambridge MA: MIT Press.

Clark A (1996) Being There. Putting Brain, Body, and World Together Again. Cambridge MA: MIT Press.

Clark RE, Squire LR (1998) Classical conditioning and bain systems: The role of awareness. Science 280: 77-81.

Freeman WJ [1975] Mass Action in the Nervous System. New York: Academic Press.

Freeman WJ [1992] Tutorial in Neurobiology. International Journal of Bifurcation and Chaos 2: 451-482.

Freeman WJ [1999] Comparison of brain models for active versus passive perception. Journal of Information Sciences 116: 97-107.

Freeman WJ [2000] Neurodynamics. An Exploration of Mesoscopic Brain Dynamics. London: Springer-Verlag.

Freeman WJ [2001] How Brains Make Up Their Minds. New York: Columbia University Press.

Haken H (1983) Synergetics: An Introduction. Berlin: Springer.

Hendriks-Jansen, Horst. (1996) Catching ourselves in the act: Situated activity, interactive emergence, evolution, and human thought. Cambridge, MA: MIT Press.

Herrick CJ (1948) The Brain of the Tiger Salamander. Chicago IL: University of Chicago Press.

Kay LM, Lancaster L, Freeman WJ (1996) Reafference and attractors in the olfactory system during odor recognition. International Journal of Neural Systems 7: 489-496.

Kay LM, Freeman WJ (1998) Bidirectional processing in the olfactory-limbic axis during olfactory behavior. Behavioral Neuroscience 112: 541-553.

O'Keefe J & Nadel L (1978) The Hippocampus as a Cognitive Map. Oxford UK: Clarendon.