

The Neurodynamics of Intentionality in Animal Brains May Provide a Basis for Constructing Devices that are Capable of Intelligent Behavior

Walter J Freeman

Department of Molecular and Cell Biology
University of California, Berkeley CA 94720-3200 USA

Essay prepared for a "Workshop on Metrics for Intelligence" in a program for "Development of Criteria for Machine Intelligence" at the National Institute of Standards and Technology (NIST), Gaithersburg MD, 14-16 August 2000

ABSTRACT

Intelligent behavior is characterized by flexible and creative pursuit of endogenously defined goals. It has emerged in humans through the stages of evolution that are manifested in the brains and behaviors of other animals. Intentionality is a key concept by which to link brain dynamics to goal-directed behavior. The archetypal form of intentional behavior is an act of observation through time and space, by which information is sought for the guidance of future action. Sequences of such acts constitute the key desired property of free-roving, semi-autonomous devices capable of exploring remote environments that are inhospitable for humans. Intentionality consists of (a) the neurodynamics by which images are created of future states as goals; (b) command sequences by which to act in pursuit of goals; (c) prediction of changes in sensory input resulting from intended actions (reafference); (d) evaluation of performance; and (e) modification of the device by itself in learning from the consequences of its intended actions. These principles are well known among psychologists and philosophers. What is new is the development of nonlinear mesoscopic brain dynamics, by which to apply chaos theory in order to understand and simulate the construction of meaningful patterns of endogenous activity that implement the perceptual process of observation. The prototypic hardware realization of intelligent behavior is already apparent in certain classes of robots. The chaotic neurodynamics of sensory cortices in pattern recognition is ready for hardware embodiments, which are needed to provide the eyes, noses and ears of devices for survival and autonomous operation in complex and unpredictable environments.

Key Words: *Chaos theory, Intentionality, Mesoscopic Brain dynamics, Perception, Reafference*

1.0 Neurodynamics of intentionality in the behavioral act of observation

1.1 *The properties of intentionality*

The first step in pursuit of an understanding of intentionality is to ask, what happens in brains 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 and interprets the resulting barrage of sensory stimuli. The concept of intentionality has been used to describe this process in different contexts, since its first use by Aquinas in 1272 [1]. The three salient characteristics of intentionality as it was developed by him are (a) intent or directedness toward some future state or goal, (b) wholeness, and (c) unity [12]. These three aspects correspond to current 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, also known as the symbol-grounding problem].

Intent comprises the endogenous initiation, construction, and direction of behavior into the world. It emerges from brains. Humans, animals and autonomous robots select their own goals, plan their own tactics, and choose when to begin, modify, and stop sequences of action. Humans at least are subjectively aware of themselves acting, but consciousness is not a necessary property of intention. Unity appears in the combining of input from all sensory modalities into *Gestalts*, 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 and emotion, but again this is not intrinsic to intention.

Wholeness is revealed by the orderly changes in the self and its behavior that constitute the development, maturation and adaptation of the self, within the constraints of its genes or design principles, and its material, social and industrial environments. Subjectively, wholeness is revealed in the remembrance of self through a lifetime of change, although the influences of accumulated and integrated experience on current behavior are not dependent on recollection and recognition. In brief, simulation of intentionality should be directed toward replicating the mechanisms by which goal states are constructed, approached and evaluated, and not toward emulating processes of consciousness, awareness, emotion, etc. in machines.

1.2 The limbic system is the chief 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 [17], and in rats with neocortex chemically inactivated by spreading depression [3]. Intentional behavior is severely altered or absent after major damage to the medial temporal lobe of the basal forebrain, as manifested most widely in Alzheimer's disease.

Phylogenetic evidence comes from observing intentional behavior in salamanders, which have the simplest of the existing vertebrate forebrains [21, 28]. The three main parts are sensory (which, as in small mammals, is predominantly olfactory), motor, and associational (Figure 1). These parts can be judged to comprise the limbic system in all vertebrates, but in the salamander they have virtually none of the "add-ons" found in brains of higher vertebrates, hence the simplicity. The associational part contains the primordial hippocampus with its interconnected septum and amygdaloid nuclei, striatal nuclei, which are identified in higher vertebrates as the locus of the functions of spatial orientation (the "cognitive map") and temporal integration in learning (the

organization of long and short term memory). These processes are essential, inasmuch as intentional action takes place into the world, and even the simplest action, such as searching for food or evading predators, requires an animal to know where it is with respect to its world, where its prey or refuge is, and what its spatial and temporal progress is during sequences of attack or escape. The feedback loops that support the flow of neural activity in the neurodynamics of intentionality are schematized in Figure 2.

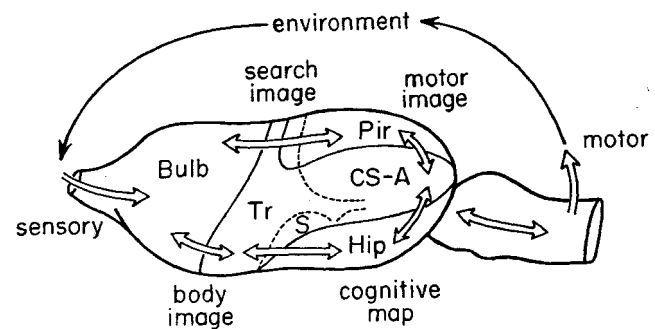


Figure 1. This schematic illustrates the sensory, motor, and associational components of the right hemisphere (seen from above) of the simplest extant vertebrate brain in the salamander. The bidirectional connections between these 3 major subdivisions of the forebrain provide for the macroscopic interactions that support the neurodynamics of the process of intentionality: goal formation, action, perception, and learning from the sensory consequences of the action taken into the environment. These components are form the prototype of the limbic system, which is found in all vertebrate brains, typically buried within exuberant growth of other "add-on" structures that operate in concert with the limbic system.

1.3 Neurodynamic manifestations of intentionality in brain activity of the primary sensory cortices: the EEG (electroencephalogram, 'local field potential')

The crucial question for neuroscientists is, how are the patterns of neural activity that sustain intentional behavior constructed in brains prior to perception? An answer is provided by studies of electrical activity of the primary sensory cortices of animals that trained to respond to conditioned

stimuli [2, 8, 10-12, 14, 22, 23]. The construction is not by recall of stored patterns but by pattern formation in distributed nonlinear systems with connections that have been modified cumulatively through learning. The manner in which this take place involves hierarchical ordering of neural activity between microscopic, mesoscopic and macroscopic levels having differing time and space scales. Cortical neurons are selectively activated by sensory receptors and made to generate microscopic activity in the form of trains of action potentials (pulses) on their axons. These and neighboring neurons by their synaptic interactions form a population forms that "binds" their activity into mesoscopic patterns [14, 18, 19, 29, 30]. These mesoscopic brain activity patterns are revealed by electrical fields of potential (EEGs) generated by interactive masses of neurons are induced by the arrival of stimuli, which trigger sequences of 1st order state transitions. These sequential states in turn converge into integrated macroscopic patterns that occupy the entirety of each cerebral hemisphere and give rise to the global patterns of brain activity, that may be related to the patterns of metabolic activity that are revealed by non-invasive brain imaging (fMRI, PET, SPECT, etc.).

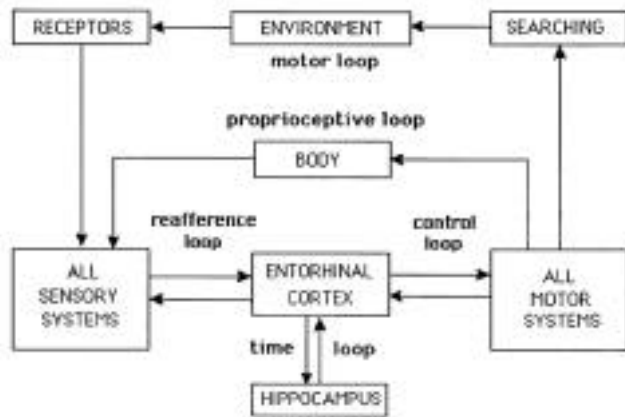


Figure 2. This diagram of brain state space maps the multiple feedback loops that support the intentional arc.

Flow of neural activity inside the brain is in two directions. Forward flow from the sensory systems to the entorhinal cortex and on to the motor systems is by spatial AM patterns of action potentials at the **microscopic** level, by which

transmitting cortices drive the neurons in their targets. Feedback flow from the motor systems to the entorhinal cortex by control loops, and from the entorhinal cortex to the sensory systems inside the brain, is by spatial AM patterns of action potentials at the **mesoscopic** level. This feedback constrains and modulates the microscopic activity in the forwardly transmitting populations. The mesoscopic feedback messages are order parameters that bias the attractor landscapes of the sensory cortices in preafferece. Forward flow supports motor output and provides the content of percepts. Feedback flow supports integrative processes in learning that lead to the wholeness of intentionality. They enable the formation of a **macroscopic** AM pattern that reflects the integration of the activity of an entire hemisphere.

Owing to the nonlinear state transitions by which they form, these mesoscopic brain states are not representations of stimuli, nor are they simple effects caused by stimuli. Each learned stimulus serves to elicit the construction of a pattern that is shaped by the synaptic modifications between cortical neurons from prior learning, which vastly outnumber the synapses formed by incoming sensory axons, and also by the brain stem nuclei that bathe the forebrain in neuromodulatory chemicals. Each cortical activity pattern is a dynamic operator that creates and carries the meanings of stimuli for the recipient 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, which together comprise the neural process of "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 provides a neural matrix of interconnections, that serves to generate continually the neural activity that forms goals and directs behavior toward them.

EEG evidence shows that the process occurs in discontinuous steps, like frames in a motion picture. Each step follows a dynamic state transition, in which a complex 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 that emerge, spread into other parts of the brain, and then dissolve to give rise to new ones, a process that Japanese mathematicians have described as "chaotic itinerancy" between "attractor ruins" [34]. Its output controls the brain stem nuclei that serve to regulate its own 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. It may be said that the neurodynamics of the limbic system, assisted by other parts of the forebrain such as the frontal lobes, initiates the novel and creative behavior seen in search by trial and error.

The limbic activity patterns of directed arousal and search are sent into the motor systems of the brain stem and spinal cord. Simultaneously, patterns are transmitted to the primary sensory cortices, preparing them for the consequences of motor actions. This process has been called "reafference" [12, 35], "corollary discharge" [32], "focused arousal" [29] and "preafference" [22, 23]. It sensitizes sensory systems to anticipated stimuli prior to their expected times of arrival. Sensory cortical constructs consist of brief staccato messages to the limbic system, which convey what is sought and the result of the search. After multisensory convergence, the spatiotemporal activity pattern in the limbic system is up-dated through temporal integration in the hippocampus. Between sensory messages there are return updates from the limbic system to the sensory cortices, whereby each cortex receives input that has been integrated with the output of the others, reflecting the unity of intentionality. Everything that a human or an animal knows comes from this iterative circular process of action, reafference, perception, and up-date. It is done by successive frames that involve repeated state transitions and self-organized constructs in the sensory and limbic cortices. This neurodynamic system is defined here as the "limbic self" in the brain of an individual, where intentional behavior is created, with help from other parts of the forebrain.

An act of observation comprises Aquinas' intentional action of "stretching forth" and learning from the consequences. It embodies the existential "action-perception cycle" of Merleau-Ponty [26]. It corresponds to Piaget's [27] cycle of "action, assimilation, and adaptation" in the sensorimotor stage of childhood development. His postulated sequences of equilibrium, disequilibrium, and re-equilibration conform to state transitions in brain dynamics, which initiate and sustain action, construct dynamic patterns in the sensory cortices, and up-date the limbic patterns by modifying synapses in the learning that follows the sensory consequences of intended actions. For Piaget, cause and effect are chains of events that have the appearance of linkage corresponding to the unfolding experience of that exploration, by which a child is trying to make sense of its world by manipulating objects in it. The origin of causal inference is buried deeply in the pre-linguistic exploratory experience of each of us. It is not easily accessed by cognitive analysis or introspection.

We are all aware of our acts of observation. It is partly by expectation of what we are looking for through reafference, partly by perceiving the changes that our actions make in the dispositions of our bodies through proprioception, and partly by our selection of stimuli from the environment through exteroception. We perceive our intentional acts as the "causes" of changes in our perceptions, and the subsequent changes in our bodies as "effects" [12]. If this hypothesis of limbic dynamics is correct, then everything that we know we have learned through the action-perception cycle, including the iterative state changes by which it is produced in brains of animals and humans. It is this cycle, in prototypic form without need for appeal to consciousness, that must be simulated in our attempts to devise intelligent machines.

2.0 Characteristics of brain states as they are revealed by EEGs

The "state" of the brain is a description of what it is doing in some specified time period. A state transition occurs when the brain changes and does something else. For example, locomotion is a state, within which walking is a rhythmic pattern of activity that involves large parts of the brain, spinal cord, muscles and bones. The entire

neuromuscular system changes almost instantly with the transition to a pattern of jogging or running. Similarly, a sleeping state can be taken as a whole, or divided into a sequence of slow wave and REM stages. Transit to a waking state can occur in a fraction of a second, whereby the entire brain and body shift gears, so to speak. The state of a neuron can be described as active and firing or as silent, with sudden changes in the firing manifesting state transitions. Populations of neurons also have a range of states, such as slow wave, fast activity, seizure, or silence. The mathematics of nonlinear dynamics is designed to study these states and the transitions by which they are accessed and abandoned.

2.1 The problem of stability of cortical states

The most critical question to ask about a state is its degree of stability or resistance to change. Evaluation is done by perturbing an object or a system [8]. For example, an object like an egg on a flat surface is unstable, but a coffee mug is stable. A person standing on a moving bus and holding on to a railing is stable, but someone walking in the aisle is not. If a person regains his chosen posture after each perturbation, no matter in which direction the displacement occurred, that state is regarded as stable, and it is said to be governed by an attractor. This is a metaphor to say that the system goes ("is attracted") to the state through an interim state of transience. The range of displacement from which recovery can occur defines the basin of attraction, in analogy to a ball rolling to the bottom of a bowl. If the perturbation is so strong that it causes concussion or a broken leg, and the person cannot stand up again, then the system has been placed outside the basin of attraction, and a new state supervenes with its own attractor and basin.

Stability is always relative to the time duration of observation and the criteria for what is chosen to be observed. In the perspective of a lifetime, brains appear to be highly stable, in their numbers of neurons, their architectures and major patterns of connection, and in the patterns of behavior they produce, including the character and identity of the individual that can be recognized and followed for many years. Brains undergo repeated transitions from waking to sleeping and back again, coming up refreshed with a good night or irritable with insomnia, but still, giving the same persons as the night before. Personal identity is usually quite

stable. But in the perspective of the short term, brains are highly unstable. Thoughts go fleeting through awareness, and the face and body twitch with the passing of emotions. Glimpses of their internal states of neural activity reveal patterns that are more like hurricanes than the orderly march of symbols in a computer. Brain states and the states of populations of neurons that interact to give brain function, are highly irregular in spatial form and time course. They emerge, persist for a small fraction of a second, then disappear and are replaced by other states. It is the flexibility and creativeness of this process that makes it so successful in animals for their adaptation to rapidly changing and unpredictable environments, and that makes it the desired platform on which to base the design of intelligent machines.

2.2 Three types of stable cortical states

In using dynamics we approach the problem by defining three kinds of stable state, each with its type of attractor. The simplest is the point attractor. The system is at rest unless perturbed, and it returns to rest when allowed to do so. As it relaxes to rest, it has the history of what happened, but that history is lost after convergence to rest. Examples of point attractors are silent neurons or neural populations that have been isolated from the brain, and also the brain that is depressed into inactivity by injury or a strong anesthetic, to the point where the EEG has gone flat (Figure 3, bottom trace). 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. 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 in 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, clapping and dancing, or it is pathological, as in the periodic oscillations of the eyes in nystagmus, or of the limbs during Parkinsonian tremor, or of the cortex during the hypersynchrony of partial complex seizures that are revealed by near-periodic spike trains (Figure 3, top trace).

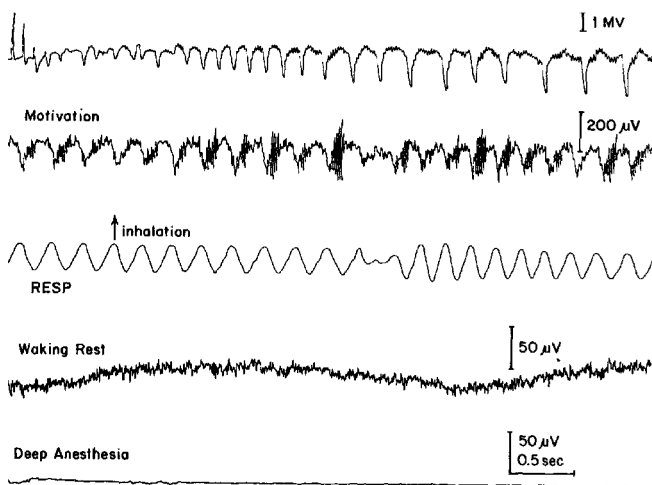


Figure 3. Four levels of function of the olfactory system are revealed by EEG recording.

The lowest is the non-interactive 'open loop' state imposed by deep anesthesia, which suppresses brain activity. The next is the resting steady state with broad spectrum $1/f^2$ aperiodic waves. The aroused level in which behavior is generated is shown by the repeated state transitions, by which bursts are formed that reveal spatial patterns of AM (amplitude modulation) relating to odorant recognition with inhalation. The upper trace shows the pattern of high-amplitude spikes when an epileptic seizure has been triggered by powerful electrical stimulation. This state is likewise chaotic, but with a reduced correlation dimension. This state also occurs during recovery from deep anesthesia on the way to the resting state [9, 31].

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 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 systems with chaotic attractors have a small number of components and a few degrees of freedom, as for example, the double-hinged pendulum, the dripping faucet, and the Lorenz, Chua, and Rössler attractors [13]. These simple models are stationary, autonomous, and noise-free, forming the class of "deterministic chaos". Large and complex real-world systems, which include neurons and neural populations are noisy, infinite-dimensional, nonstationary, non-autonomous, yet capable of chaotic behavior which has been called "stochastic chaos" [14]. The source is postulated to be the synaptic interaction of millions of neurons, which create fields of microscopic noise in cortex, but which are constrained by their own interactions to generate mesoscopic order parameters that regulate the spatiotemporal patterns of cortical activity revealed by the EEG. These spatiotemporal patterns are revealed by spatial patterns of amplitude modulation ("AM patterns") of a spatially coherent aperiodic carrier wave in the gamma range of the EEG. They appear in time series as bursts of oscillation (Figure 3), and their spatial patterning indicates the existence of an attractor landscape, which is actualized in the olfactory system with each inhalation (Figures 4 and 5 during intentional behavior).

The discovery that brain dynamics operates in chaotic domains has profound implications for the study of higher brain function [31]. 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 collection of attractors, each with its basin, and that it can move from one to another in an itinerant trajectory [34].

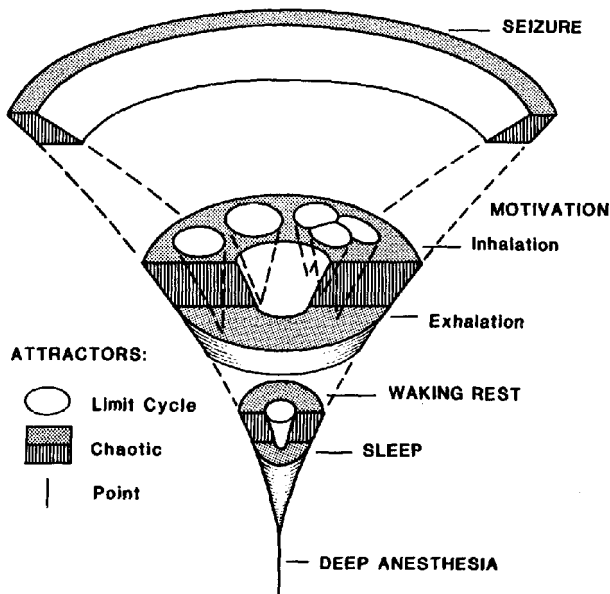


Figure 4. A bifurcation diagram of the olfactory system state space is constructed from the EEGs in Figure 3.

It retains in its pathway across its basins its history, which fades into its past, just as its predictability into its future decreases. Transitions between chaotic states constitute the dynamics that we need to understand how brains perform such remarkable feats as abstraction of the essentials of figures from complex, unknown and unpredictable backgrounds, generalization over examples of recurring objects never twice appearing the same, reliable assignment to classes that lead to appropriate actions, and constant up-dating by learning.

2.3 *The 1st order cortical state transition is an elemental step in intention*

Systems such as neurons and brains that have multiple chaotic attractors also have point and limit attractors, each with its basin of attraction, which serves to provide the generalization gradient required for perception of recurring stimuli that are never twice the same. If the basin is that of a point or a limit cycle attractor, the system can proceed predictably to an identical end state. If the basin leads to a chaotic attractor, the system goes into ceaseless fluctuation, as long as its energy lasts.

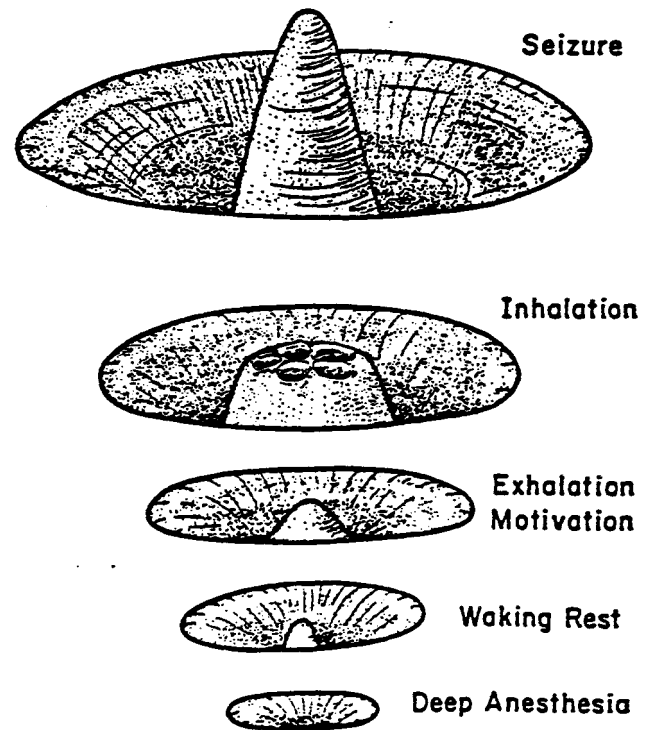


Figure 5. This perspective drawing of a projection from an infinite dimensional brain state space into 3-space offers a view of how an attractor landscape of learned basins of attraction is created with each inhalation. The selection is made by the input odorant. If the stimulus is novel or unknown, the system goes into the chaotic well, which provides the aperiodic unpatterned activity that drives Hebbian learning for new basin formation. From Skarda and Freeman 1987.

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. 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. A deterministic chaotic system that is in the basin of one of its chaotic attractors is legendary for its sensitivity to the initial conditions. 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 has the capacity to create novel patterns

constituting endogenous increases in information in the course of continually constructing its own trajectory into the future.

Our EEG evidence indicates that every primary sensory cortex maintains multiple basins corresponding to previously learned classes of stimuli, as well as to the unstimulated state, which together form an attractor landscape. They all show evidence that the vehicle they use for transmission of their output is an aperiodic carrier wave that is amplitude-modulated in the two spatial dimensions of cortical coding, and that is gated by extra-cortical forcing functions in the theta range (2-7 Hz). We note that we predicted a common code for all sensory systems, on the basis that the signals from all sensory cortices must be combined in the limbic system to form gestalts. We postulate that preafferent input from the limbic system can serve to bias the landscapes in such a way as to facilitate the capture of the multiple sensory systems by basins of the attractors corresponding to the goal of the intended observation, perhaps in the manner of the variable tiling in a Voronoi diagram. This chaotic prestimulus state of expectancy establishes the sensitivities of the cortices, so that the very small number of sensory action potentials evoked by the expected stimuli can simultaneously carry the cortical trajectories into the basins of the appropriate attractors as they are created by the forcing function, in the case of olfaction by inhalation (Figure 5), irrespective of which equivalent receptors actually receive the expected stimuli in the different sensory modalities. In the absence of the stimulus, the cortices continue to transmit their outputs to the limbic system, confirming the continuing absence. The stimuli are also selected by the limbic system through orientation of the sensory receptors in space by sniffing, looking, and listening. We believe that the basins of attraction in each of the sensory cortices are shaped by limbic input to sensitize them for receiving and processing the desired class of stimuli in every modality, whatever may be the goal at the moment of choice.

3.0 Problems in use of chaotic dynamics in the development of advanced machine intelligence

Chaotic dynamics has proved to be extremely difficult to harness in the service of intelligent machines. Most studies that purport to control chaos either find ways to suppress it and replace it

with periodic or quasiperiodic fluctuations, or to lock two or more oscillators into synchrony sharing a common aperiodic wave form, often as an optimal means for encryption and secure transmission. Our aim is to employ chaotic dynamics as the means for creating novel and endogenous space-time patterns, which must be the means to achieve any significant degree of autonomy in devices that must operate far from human guidance, where in order to function they must make up their courses of action as they go along. We know of no other way to approach a solution to the problem of how to introduce creative processes into machines, other than to simulate the dynamics we have found in animal brains. To be sure, there are major unsolved problems in this approach, chief among them that we know too little about the dynamics of the limbic system. Hence we find it necessary to restrict the development of hardware models to the stage of brain-world interaction that we know best, which is the field of perception. In brief, what are the problems in giving eyes, ears and a nose to a robot, so that it might learn about its environment in something like the way that even the simpler animals do - by creating hypotheses and testing them through their own actions?

3.1 Noise stabilization of chaotic dynamics, opening the way to analog-digital hybrid embodiments

The operations in the olfactory system by which the state transitions and pattern constructions for pattern classification are simulated in software and hardware embodiments have been described in a series of publications [9-12, 14]. Our simulations are done with a set of ~920 interconnected 1st-order nonlinear ordinary differential equations [Figure 6], forming what we have named the KIII model [8]. The basic element, the KO set, is a 2-stage linear integrator simulated in hardware [6, 7] by 2 operational amplifiers, whose output is passed through an asymmetric sigmoid function modeled by 2 diodes back-to-back. Connections between 64 elements are time multiplexed (Figure 7) through a MUX, an amplifier with voltage-controlled gain, and a DMUX [10]. Switching is controlled by a digital computer at a clock rate suitable for the pass band of the carrier wave. For each connected pair the gain is stored in memory, so that the connection strengths are easily modified during learning. With this device the connectivity grows by 2-N instead

of by N^2 . In digital embodiment the equations have been solved by numerical integration on Unix, Macintosh, and PC platforms, and by vector programming on the Cray M/X.

Interaction of KO sets of like kind (excitatory or inhibitory) giving point attractors is modeled by KI sets; interaction of KIe and KIi sets giving limit cycle attractors is modeled by KII sets. Three serial KII sets in layers that correspond to the olfactory bulb, prepyriform cortex, and an intervening control nucleus called the AON is modeled by the KIII set; if the 3 characteristic frequencies are incommensurate, and the feedback delays between the 3 layers are distributed to act as low-pass filters, the solutions of the equations give the aperiodic waveforms and broad $1/f^2$ spectra (Figure 8) of EEGs from the 3 layers. The asymmetric sigmoid endows the system with the property of nonlinear state transitions on step inputs, owing to the amplitude-dependent gain of the KO elements.

In the course of digital simulation it has become apparent that a minimum of 64 elements will suffice for 2-D pattern classification under Hebbian and non-Hebbian reinforcement learning [16, 24, 25, 37, 38]. The large number of equations leads to attractor crowding [15], in which the basins of attraction shrink close to the size of the digitizing step in using rational numbers for computation, so that sooner or later the system jumps out of its designated chaotic basin into a neighboring basin that is most likely to be that of a point or limit cycle attractor, which kills the system. This problem has been solved by use of additive noise on the order of 15% of the amplitude of the aperiodic state variables [4, 5, 13, 15], giving robust attractor landscapes for learning and pattern classification [24, 25]. The lesson learned is that deterministic chaos, in which the system is low-dimensional, stationary, strictly autonomous, and noise-free, is inappropriate for modeling biological and machine intelligence. Brains operate with what we call 'stochastic chaos' [13], which is high-dimensional, nonstationary with regularly repeated state transitions, engaged with its surround, and deeply embedded in noise created by KIe sets and manifested in high densities of action potentials. The noise in digital models is simulated with random number generators, either rectified to simulate KIe sets or off-set with d.c. bias to simulate the noise in KIIIe sets.

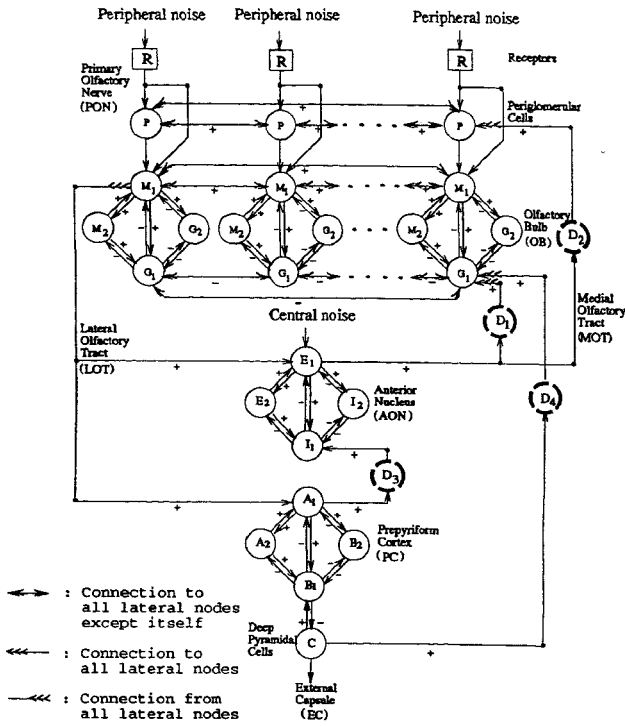


Figure 6. Schematic of KIII model.

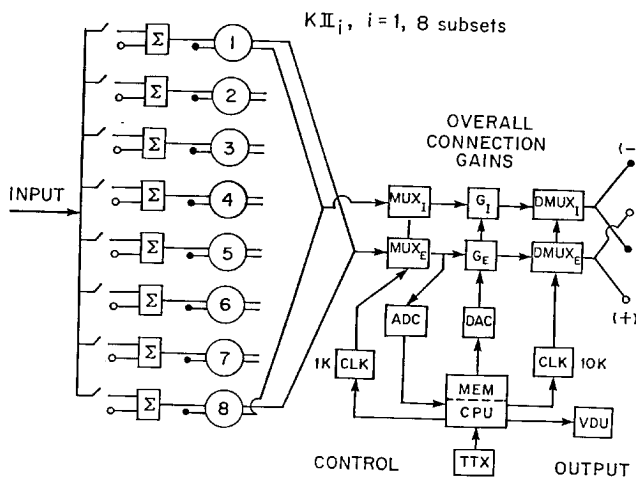


Figure 7. Schematic for connecting KII sets by multiplexing.

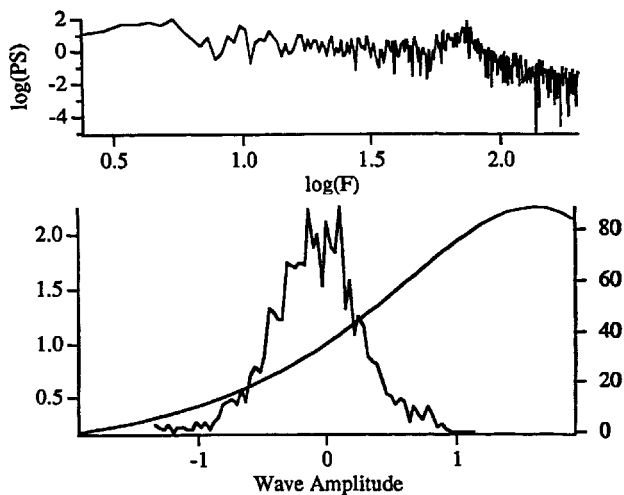


Figure 8. The power spectrum and amplitude histogram for a simulated EEG trace from the KIII model, with a section of the asymmetric nonlinear gain curve, showing the nature of the nonlinearity that provides for destabilization by the input. The interactive gain increases with excitatory input. From Freeman et al. 1997.

The finding in digital embodiments that noise is not only unavoidable but is necessary for stable high-dimensional chaotic dynamics opens the way to analog embodiments [7], in which noisy components resemble the characteristics of local pools in nerve cell assemblies, but which offer much higher rate of temporal and spatial integration, the use of continuous variables in place of rational numbers, and the feasibility of implementing the dynamics on chips suitable for incorporation into mobile devices.

3.2 Embedding devices for perception into autonomous cognitive machines

The KII sets have multiple robust limit cycle attractors, which become embedded as chaotic attractors when coupled in serial layers with distributed delayed feedback. The KIII model is offered as the prototype for constructing devices in hardware and software to implement the elementary steps of perception, thus providing robots with the sensory ports that they need to guide them through their environments. These steps are the interpretive operations necessary to normalize, compress, abstract and generalize over successive inputs preparatory to classification [5,

16, 25, 33, 36]. These cognitive operations are done by the nonlinear operations in the input stage and by the basins of attraction in the landscape formed by learning in each of the sensory systems. They are required in each of the ports providing information to the mobile device about its visual, auditory, tactile, and chemical environments. Our tests of the KIII model have shown that it can learn a new class in half a dozen trials instead of the thousands of trials required by MLPs, and that new learning occurs without degradation of previous attractors, although, as in the case of the olfactory system, the attractors are modified through attractor crowding. The superior level of 'intelligence' is demonstrated by the capacity of the KIII model to separate items in 64-space that belong to identifiable classes but are not linearly separable. The classes are, in fact, constructed by the model and are not imposed from outside, constituting an aspect of autonomy. In other words, the system creates its own features from its own experience of the constancy of relations between channels in the 8x8 64-channel input array.

Formation of a world-view by which the device can guide its explorations for the means to reach its goals depends on the integration of the outputs of the several sensory systems, in order to form a multisensory percept known as a gestalt. This integration is easily done when all of the ports have their outputs in the same form: a vector consisting of a 2-D spatial pattern of amplitude modulation of a 1-D aperiodic wave form in the gamma range (nominally 30-60 Hz), which is segmented in time at a frame rate of nominally 2-7 Hz and frame durations on the order of 0.1 sec. Precise clocking and synchronization are not prerequisite.

The sequential frames deriving from sampling the environment must then be integrated over time and oriented in space. An example of how these higher operations might be done was provided by W. Gray Walter [36] with his electronic tortoises, which had the capacity for autonomous goal-directed search involving the adjudication of conflicting needs in an uncertain environment.

The performances of these devices set a challenging level of 'intelligence' to which to aspire, and they also serve to highlight some of the difficulties in using the descriptive term "autonomous". As with animals the devices were

untethered, and they learned to avoid obstacles without need for instruction or intervention, if within their limited capacities for locomotion. However, they were programmed to satisfy their own needs without regard for or comprehension of anything else's, perhaps in analogy to house pets, whose sole purpose, however inadvertent, is to provide enjoyment to their owners, and seldom to do useful work or bend their talents to the benefits of the owners, or, in the case of the machines, the designers and builders.

It is already apparent that fully autonomous vehicles are not in the best interest of researchers and the general public, except as demonstrations of what might emerge as major problems from this line of study. It is also clear that such devices can and will be built, and that the proper path of future management will not be by techniques of training and aversive conditioning, but by education, with inculcation of desired values determined by the manufacturers that will govern the choices that must by definition be made by the newly autonomous mechanical devices.

Acknowledgments

This work was supported by grants from the National Institute of Mental Health MH06686, the Office of Naval Research N00014-90-J-4054, and from ARO MURI DAAH04-96-1-0341.

References

- 1] Aquinas, St. Thomas. (1272). *Treatise on Man*. In: Summa Theologica. Translated by Fathers of the English Dominican Province. Revised by Daniel J Sullivan. Published by William Benton as Volume 19 in the Great Books Series. Chicago: Encyclopedia Britannica, Inc., 1952.
- 2] 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.
- 3] Bures J, Buresová O, Krivánek J (1974) *The Mechanism and Applications of Leão's Spreading Depression of Electroencephalographic Activity*. New York: Academic Press.
- 4] Chang H-J, Freeman WJ (1998) Biologically modeled noise stabilizing neurodynamics for pattern recognition. *International Journal of Bifurcation & Chaos* 8 (2), 321-345.
- 5] Chang H-J, Freeman WJ (1999) Local homeostasis stabilizes a model of the olfactory system globally in respect to perturbations by input during pattern classification. *International Journal of Bifurcation and Chaos* 8: 2107-2123.
- 6] Edelman JA, Freeman WJ (1990) Simulation and analysis of a model of mitral-granule cell population interactions in the mammalian olfactory bulb. *Proceedings IJCNN I*: 62-65.
- 7] Eisenberg J, Freeman WJ, Burke B (1989) Hardware architecture of a neural network model simulating pattern recognition by the olfactory bulb. *Neural Networks* 2: 315-325.
- 8] Freeman WJ (1975) *Mass Action in the Nervous System*. New York: Academic Press.
- 9] Freeman WJ (1987) Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biological Cybernetics* 56: 139-150.
- 10] Freeman WJ (1988) Pattern learning and recognition device. United States Patent # 4, 748, 674, May 31, 1988.
- 11] Freeman WJ (1992) Neurons to Brain Chaos. *International Journal of Bifurcation and Chaos* 2: 451-482.
- 12] Freeman WJ (1995) *Societies of Brains*. Mahwah NJ, Lawrence Erlbaum Associates.
- 13] Freeman WJ (1999) Noise-induced first-order phase transitions in chaotic brain activity. *International Journal of Bifurcation and Chaos* 9: 2215-2218.
- 14] Freeman WJ [2000] *Neurodynamics. An Exploration of Mesoscopic Brain Dynamics*. London UK: Springer-Verlag.
- 15] Freeman WJ, Chang H-J, Burke BC, Rose PA, Badler J (1997) Taming chaos: Stabilization of aperiodic attractors by noise.

- IEEE Transactions on Circuits and Systems 44: 989-996.
- 16] Freeman WJ, Yao Y, Burke B. (1988) Central pattern generating and recognizing in olfactory bulb: A correlation learning rule. *Neural Networks* 1: 277-288.
- 17] Goltz FL (1892) Der Hund ohne Grosshirn. Siebente Abhandlung über die Verrichtungen des Grosshirns. *Pflügers Archiv* 51: 570-614.
- 18] Gray CM (1994) Synchronous oscillations in neuronal systems: mechanisms and functions. *Journal of Comparative Neuroscience* 1: 11-38.
- 19] Haken H (1983) *Synergetics: An Introduction*. Berlin: Springer.-Verlag
- 20] Hardcastle VG (1994) Psychology's binding problem and possible neurobiological solutions. *Journal of Consciousness Studies* 1: 66-90.
- 21] Herrick CJ (1948) *The Brain of the Tiger Salamander*. Chicago IL: University of Chicago Press.
- 22] Kay LM, Freeman WJ (1998) Bidirectional processing in the olfactory-limbic axis during olfactory behavior. *Behavioral Neuroscience* 112: 541-553.
- 23] Kay LM, Lancaster L, Freeman WJ (1996) Refference and attractors in the olfactory system during odor recognition. *International Journal of Neural Systems* 7: 489-496.
- 24] Kozma R, Freeman WJ (1999) A possible mechanism for intermittent oscillations in the KIII model of dynamic memories - the case study of olfaction. *Proceedings, IJCNN'1999, Washington DC*.
- 25] Kozma R, Freeman WJ (2000) Encoding and recall of noisy data as chaotic spatio-temporal memory patterns in the style of the brains. *Proceedings, ICJNN'2000. Como, Italy*
- 26] Merleau-Ponty M (1942) *The Structure of Behavior* (AL Fischer, Trans.). Boston: Beacon Press (1963).
- 27] Piaget J (1930) *The child's conception of physical causality*. New York: Harcourt, Brace. p. 269
- 28] Roth G (1987) *Visual Behavior in Salamanders*. Berlin: Springer-Verlag
- 29] Sheer DE (1989) Sensory and cognitive 40-Hz event-related potentials: Behavioral correlates, brain function, and clinical application. *Brain Dynamics*. Basar E, Bullock TH (eds.) Berlin: Springer-Verlag.
- 30] Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience* 18: 555-586.
- 31] Skarda CA, Freeman WJ (1987) How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences* 10: 161-195.
- 32] Sperry RW (1950) Neural basis of the spontaneous optokinetic response. *Journal of Comparative Physiology* 43: 482-489.
- 33] Storm C, Freeman WJ (1999) A novel dynamical invariant measure addresses the stability of the chaotic KIII neural network. *Proceedings, IJCNN'1999, Washington DC*.
- 34] Tsuda I (1996) A new type of self-organization associated with chaotic dynamics in neural networks. *International Journal of Neural Systems* 7: 451-459.
- 35] von Holst E & Mittelstaedt H (1950) Das Refferenzprinzip *Naturwissenschaften* 37: 464-476.
- 36] Walter WG (1963) *The Living Brain*. New York: Norton.
- 37] Yao Y, Freeman WJ (1990) Model of biological pattern recognition with spatially chaotic dynamics. *Neural Networks* 3: 153-170.
- 38] Yao, Y., Freeman WJ, Burke, B., Yang, Q. (1991) Pattern recognition by a distributed neural network: An industrial application. *Neural Networks* 4: 103-121.