SENSE, THOUGHT, AND CONSCIOUSNESS

Gerald E. Marsh

The late 19^{th} and 20^{th} centuries brought a revolution in the scientific understanding of the universe around us, one whose effects are still being felt around the world as it forces people to change their conception of the universe and the place of human beings within it. While many interesting questions remain, we now understand the evolution of the universe from the first few moments of its coming into existence some fourteen billion years ago—from the creation of matter to the formation of galaxies, stars, and planets.

Our conception of our own place in the universe continues to dramatically change. It is now known from direct observation that almost all stars have planetary systems and we can expect to soon find evidence that many harbor life forms. Advances in biology, and our understanding of the evolution of life, have also grown enormously. This new knowledge will enable us, during this century, to design life forms for various practical purposes raising many religious and ethical questions. This has already begun. Our current understanding of the physical world including its biological aspects compared to even half a century ago is simply stunning! It cannot help but inspire a sense of awe and wonder in those who are fortunate enough to come to understand it.

Perhaps the greatest remaining mystery is the nature of consciousness itself, which has been a subject of human inquiry for at least the last several millennia for which we have records.

The origin of much of the difficulty in even defining consciousness dates back to the mind-body dualism introduced by René Descartes, which effectively removed the study of consciousness from scientific inquiry. In doing so Descartes left open the possibility of identifying the mind with the soul, thereby avoiding conflict with his own beliefs and those of the Catholic Church.

Intelligent thought has long been believed to be restricted to human beings. Those with a religious orientation believe that human beings are free to choose between good and evil. Such a choice would, of course, be meaningless without the innate ability for conscious, intelligent thought. Consciousness and the ability to learn are the essence of what one means by the immortal soul, although from a strictly religious perspective the soul might be called a consequence of God's grace. But these religious beliefs would be directly challenged if it were found that consciousness and thought was a natural biological phenomenon.

More than half a century ago, Homer Smith wrote a beautiful book called *Man and His Gods*. It had a forward by Albert Einstein who characterized Smith's exposition as follows:

> "The work is a broadly conceived attempt to portray man's fear-induced animistic and mythic ideas with all their far-flung transformations and interrelations. It relates the impact of these phantasmagorias on human destiny and the causal relationships by which they have become crystallized into organized religion."

In the *Epilogue*, Smith captures what must happen if the modern world is to avoid what might well be characterized as a social form of Armageddon:

> "As a fallen angel, man would be ludicrous. As an intelligent animal, he has reason to be proud because he is the first who can ask himself, "Whither, Why, and Whence?' and confident because he can know himself as a creature of earth who has risen by his own efforts from a low estate. If he would rise higher he must be true to earth, he must accept that he is its creature, unplanned, unprotected and unfavored, co-natural with all other living creatures and with the air and water and sunlight and black soil from which their dynamic pattern has been fabricated by impersonal and indifferent forces. In every wish, thought and action he is seeking to escape the same protoplasmic disquietude that impels the meanest flesh crawling beneath his feet. He must find his values and his ends entirely within this frame of reference."

In what follows, I will argue that that consciousness and thought is indeed a natural biological phenomenon. The basis for all modes of thought used by animals including human beings can be traced back to their sensory perceptions. This is covered in the first section along with a discussion of the nature of consciousness. An exploration of vision, which plays a central role in some kinds of thought, follows. This leads to a section on primary consciousness followed by one on higher-order consciousness. Of course, the most important issue is the emergence of the "I" or inner self that underlies the Descartes duality. This is discussed in the following section. There are then sections on the ultimate source of consciousness, emergence in biological neural networks, and artificial intelligence.

Modes of Thought

Remembrance and Reflection, how ally'd;
What thin partitions Sense from Thought divide.

Alexander Pope, Essay on Man, Epistle I

Human beings have a variety of senses including the obvious ones of sight, hearing, touch, smell, and taste. Many other animals have different suites of senses, some in common with humans and some not—like the ability to sense and use electric or magnetic fields for practical purposes such as finding prey and navigation, or having extended, or at least different, visual or hearing ranges than human beings. Animal interpretation of the world around them can be very different from each other and from human perception. Local sensory neuronal networks associated with sensory neurons that detect signals from the surrounding environment often perform some processing of the received sensory data before the resulting coded information is sent on to the brain. The brain itself is organized around receiving, processing, and storing this information upon which responses are based.

The neuronal networks of the lower animals, such as jellyfish and some insects, are almost hard wired in the sense that little if any learning is required for their full behavioral repertoire to become available after they are fully formed. In the higher animals, much of this ability to process sensory information is developed during early life and is a learned process in that, while the basic architecture of the nervous system is genetically determined, its development depends on early sensory stimulation and is therefore a learned process in the broader meaning of the term.

Thus, for example, we learn to interpret visual information as shapes, and we similarly learn to interpret a sequence of sounds as music or language. This interpretational process involves memory of previous exposure to similar sensory input. A given neural network can have many different patterns of excitation since individual neurons making up the network may or may not be an active part of any particular pattern of excitation; some are excited and some are inhibited. The number of possible combinations of even a few thousand neurons is enormous. This allows lower animals to display rather complex behavior when they have only a limited number of neurons.

Recent studies in humans, using functional magnetic resonance imaging, show that the same neural representations are activated for perception and when viewing remembered visual images, and repetition improves the correlation. If we listen to a piece of music that has been previously heard, a memory of the piece is retrieved after hearing only the first few notes, and if an incorrect note is played we notice it immediately. Similar behavior is displayed by other animals: a bird that has learned about a predator's shadow at an early age does not need to carefully categorize its image at a later appearance, it reacts instantly once enough information is available to call up the earlier memory; sometimes we see an animal on the lawn, only to have it almost instantly replaced by the crumpled paper bag it really is. Such misperceptions are common and clearly show that a camera is not a good model for vision; at best, the analogy serves to describe the projection of an image by the lens of the eye on the retina. The act of seeing itself involves visual processing by the retina and brain at the moment of seeing coupled with visual memories as imbedded in the neural networks of the visual cortex of the brain.

Animals at the level of cats and dogs probably have some ability to recall memories of various images, sounds, or other sensations. It is thought that the visual memory of squirrels is highly developed since they remember very well where they have buried food (visual or chemical markers that could give an alternative explanation have not been found); some birds also hide food and come back and successfully find it at a later time. However the ability of such animals to recall memories of past sensations independent of current sensory input, and correlate and manipulate them as thought, is certainly very limited compared to humans—although, as we shall see, there are exceptions.

Even among humans the ability to recall different types of sensory input is quite variable. Some people have superb visual memories both in space and time; they are able to order them so as to dynamically visualize, for example, complex machinery and its operation. Some visualize in color, which itself can be variable since people's color vision differs, and others do not. Many cannot do either since they have very little spatial relations ability. Beethoven was surely able to recall music and "play it in his head"; after all, he was deaf when he wrote his late quartets, arguably his best works. While often being able to whistle a tune or recognize a piece of music, few people can actually "play" instruments in their head. But some can, and sometimes with great fidelity. Some can think using internal, silent words that they actually "hear" (known as sub-vocalization), others have only very limited ability to do so. Other variations may appear in the ability to recall tastes, smells, or other senses.

Such variations are often hidden in that we all will agree, for example, on the colors, although experiencing them very differently; mildly color blind people see red and green, but quite differently from those with normal color vision; music experienced by young people, able to hear a wide range of sound frequencies, differs greatly from the experience of older people with reduced hearing range listening to the same piece of music, but they rarely notice the difference in discussion of their impressions of the piece. We abstract certain qualities from sense impressions be they visual, aural, or other and give them a common name. But the commonality can be deceptive in that variations in perceptions and human thought patterns can play an important and generally unrecognized role in miscommunication between individuals and groups in addition to educational, language, and cultural differences.

While the basic architecture of the nervous system is genetically determined, the information carrying capacity of DNA is far too small to specify the enormous number of interconnections of the brain. Instead, large quantities of excess neurons are produced during early life that form the basis for later learning. Gerald Edelman has developed the concept of neural Darwinism to help explain the selection of these neurons to form functional neural networks, and thus the micro-architecture of the brain. Learning involves the formation of memories; how these are formed and stored, a subject having a long history and literature, has now been understood at the neuronal and molecular level. It is well described in Eric Kandel's memoir and scientific exposition, *In Search of Memory*.

As the complexity of the neural structure of brains increased during evolution, consciousness gradually makes its appearance. In his book *Bright Air, Brilliant Fire: On the Matter of the Mind* Edelman distinguishes two types of consciousness, primary and higher-order. Primary consciousness is "composed of phenomenal experiences such as mental images, but it is bound to a time around the measurable present, lacks concepts of self, past, and future, and lies beyond direct descriptive individual report from its own standpoint." Higher-order consciousness "is based on the occurrence of direct awareness in a human being who has language and a reportable subjective life." Edelman does allow that other animals can think, but since they lack true language, higher-order consciousness cannot "flourish in them as it does in us." Higher-order consciousness is where one is "conscious of being conscious". We will return to this distinction later in this essay as it is directly related to the emergence of the "I" or inner self.

What appears to be the case then is this: Human beings think using a variety of basic elements related to our senses be they visual images, sub-vocalized words, or other recalled and composite constructs based on the modalities of our senses. Human beings use these elements of thought to enable what is called conceptual thought, that which involves abstraction and inference. It includes, for example, inductive and deductive logic, as well as mathematics and symbolic logic in its broadest sense. These elements can be used individually, and more often in combination. How people think of the same concepts, and which elements of thought are used and how they are combined, may vary from individual to individual.

In human beings, as put by Esther Gardner and John Martin in their on-line article *Coding of Sensory Information*, “our perceptions differ qualitatively from the physical properties of stimuli because the nervous system extracts only *certain* pieces of information from each stimulus, while ignoring others, and then interprets this information in the context of the brain's intrinsic structure and previous experience . . . . Colors, tones, smells, and tastes are mental creations constructed by the brain out of sensory experience. They do not exist, as such, outside the brain.”

The elements of thought of animals including humans are ultimately related to their sense impressions and their neural cognates. Memory allows us to conceive of objects independent of immediate sensory input; and the ability to project our thoughts into the past and into the future enables us to imagine objects and possible events as they may have been (beyond the extent of memory) and project the present into the future. Memory itself is composed of at least three different types: procedural—concerned with sequential operations; episodic—remembering personal events in the sequence they occurred; and semantic, often identified with symbolic knowledge dealing with ideas, concepts and meanings. These divisions are analytic in nature, being useful for heuristic purposes, but real memory will be a synthesis of these divisions. The ability to recall information related to these categories of memory can be expected to vary widely among individuals.

What thinking may exist in non-human animals will reflect the elements of thought available to them from the neural cognates related to their often very different senses. Vision, which we now turn to, plays a special role in thought.

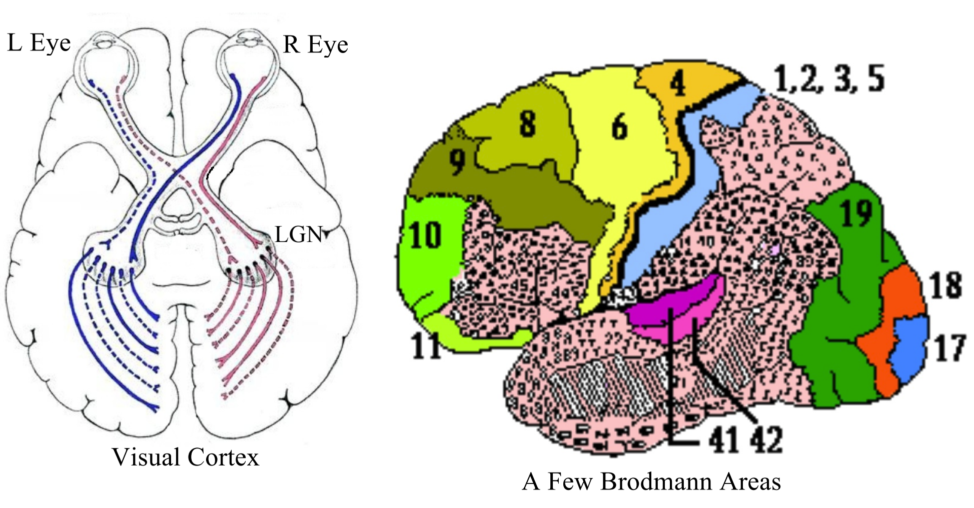
Vision

Although an analogy is often made between the eye and a camera, after an image is formed by the lens—assuming one exists—on the retina, the analogy fails abysmally.

To show this, we first consider the very primitive vision of a frog. It was elegantly described in 1968 by J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts in their article *What the Frog's Eye Tells the Frog's Brain*:

> "A frog hunts on land by vision. He escapes enemies mainly by seeing them. His eyes do not move, as do ours, to follow prey, attend suspicious events, or search for things of interest. If his body changes its position with respect to gravity or the whole visual world is rotated about him, then he shows compensatory eye movements. These movements enter his hunting and evading habits only, e.g., as he sits on a rocking lily pad. Thus his eyes are actively stabilized. He has no fovea, or region of greatest acuity in vision, upon which he must center a part of the image. He has only a single visual system, retina to colliculus†, not a double one such as ours . . . The frog does not seem to see or, at any rate, is not concerned with the detail of stationary parts of the world around him. He will starve to death surrounded by food if it is not moving. His choice of food is determined only by size and movement. He will leap to capture any object the size of an insect or worm, providing it moves like one. . . . His choice of paths in escaping enemies does not seem to be governed by anything more devious than leaping to where it is darker. . . . He does remember a moving thing providing it stays within his field of vision and he is not distracted."

Human vision is far more complicated than that of the frog and begins when the lenses of the eyes form high-resolution images on their respective retinas. There are five different types of cells in the retina including the ganglion cells whose axons exit the eye and make up the optic nerve. Information, including color, motion, and contrast travel over the



(a) (b)

Figure 1. (a) A cross section of the brain showing the partial crossing of the optic nerves and the path of the "optic tracks" to the Lateral Geniculate Nucleus (LGN); (b) some of the Brodmann areas. The primary visual cortex includes areas 17, 18, and 19.

† A small part of the frog brain involved in vision and hearing.

* Vision is believed to be mediated by three parallel pathways that process information for depth and form,

optic nerve from each eye with each nerve containing about one million axons. As is schematically illustrated in Fig. 1 (a), only half of the retinal axons from each eye cross over in the region known as the optic chiasm. In vertebrates that have eyes on the sides of their head, the optic nerves may cross entirely at the optic chiasm. In humans, where the eyes are in the front of the head so that the eyes have a large binocular overlap, we get the ~50% crossover.

The reason for this crossover is that the entire visual field, rather than the retinal field, is projected onto the visual cortex; that is, the axons from the side of the retina near the temples join the axons from the nasal side of the retina from the eye on the opposite side of the head. Before getting to the visual cortex, however, the "optic track" (the bundle of axons making up the optic nerve on the other side of the optic chiasm) first goes to the Lateral Geniculate Nucleus (LGN), whose purpose is to merge the two half images from each eye into one seamless visual field before sending the visual data to the visual cortex. The LGN is part of the thalamus, a dual-lobed structure under the cerebral cortex involved in sensory perception and motor function regulation. The LGN has several other functions including governing the level of consciousness and attention. It has six layers, two of which are involved with the movement of the image in both eyes and the remaining layers analyzing the eye images for detail and color information.

The LGN also has two-way communication with other brain regions. This constantly changing two-way, or "reentrant" connection between brain regions will play an important role in what follows. By reentrant connection is meant the massively parallel, two way, and changing signal paths between one brain region and another.

The concept of reentrance was introduced by Gerald Edelman in his theory of neuronal group selection. As put by him and Giulio Tononi in Chapter 9 of the book *Neural Correlates of Consciousness* published in 2000 and edited by Thomas Metzinger: "Reentry is a process of ongoing parallel and recursive signaling between separate brain maps along massively parallel anatomical connections, most of which are reciprocal. . . . It is reentry acting by neuronal group selection that assures the integration so essential to the creation of a scene in primary consciousness." And this is the point: the scene—in this case restricted to the visual scene—is the dynamic integration of information from different parts of the brain; it constantly changes from moment to moment and is central to the emergence of consciousness. In general, the "scene" would include sensory input from aural and other qualia, by which is meant "the collection of personal or subjective experiences, feelings and sensations that accompany awareness".

Most processing of visual information is done in the brain's visual cortex. Historically, Broadmann area 17 was known as the primary visual cortex and areas 18 and 19 as the visual association cortex. These regions are further divided into areas V1-V5. To get a sense of the complexity of the visual process one can do no better than to quote Larry Squire, et al. in their 2008 book *Fundamental Neoroscience*:

> "The demonstration of multiple visual cortical areas has been one of the important discoveries of the past quarter century in the field of sensory neurobiology. A vast expanse of cerebral cortex—greater than 50% of the total in many primate species—is involved primarily or exclusively in the processing of visual information. The extrastriate cortex [The primary visual, or *striate* cortex is Brodmann area 17] now includes areas 18 and 19, as well as large regions of the temporal and parietal lobes. . . . It is composed of some 30 subdivisions that can be distinguished by their physiology, cytoarchitecture, histochemistry, and/or connections with other areas. . . . Each of these extrastriate visual areas is thought to make unique functional contributions to visual perception and visually guided behavior. . . . Neurons at lower levels in the visual system are sensitive to isolated and specific features in visual scenes. Higher visual areas respond to very specific attributes, but these attributes are increasingly remote from the physical stimulus. Instead, they represent increasingly complex concepts, such as the motion of an extended object or the identity of a face."

Humans have trivalent color vision meaning that there are three different classes of cone cells in the retina that contain different pigments whose absorption spectra overlap but peak at different wavelengths of the visual spectrum—blue, green, and red. Color is a sensation produced in our brains and does not exist, as such, in the external world. It is thought that color vision processing in the nervous system compares and contrasts the output of the cone cells in pairs.

Edwin Land, the founder of the Polaroid Corporation, was fascinated by the process of color perception and in the 1950s proposed a model in which the output of each cone cell is normalized over the entire visual scene before being compared with each other locally to generate the perception of color. An interesting discussion of color vision has been given by Peter Gouras and is available at http://webvision.med.utah.edu/book/part-vii-color-vision/color-vision/ . This work contains an extensive set of references for the interested reader.

Primary Consciousness

As mentioned earlier, Edelman distinguishes two types of consciousness, primary and higher-order. He uses the diagram shown in Fig. 2A to illustrate the difference between the two types of consciousness. Figure 2B shows the location of the different brain regions referred to in Fig. 2A.

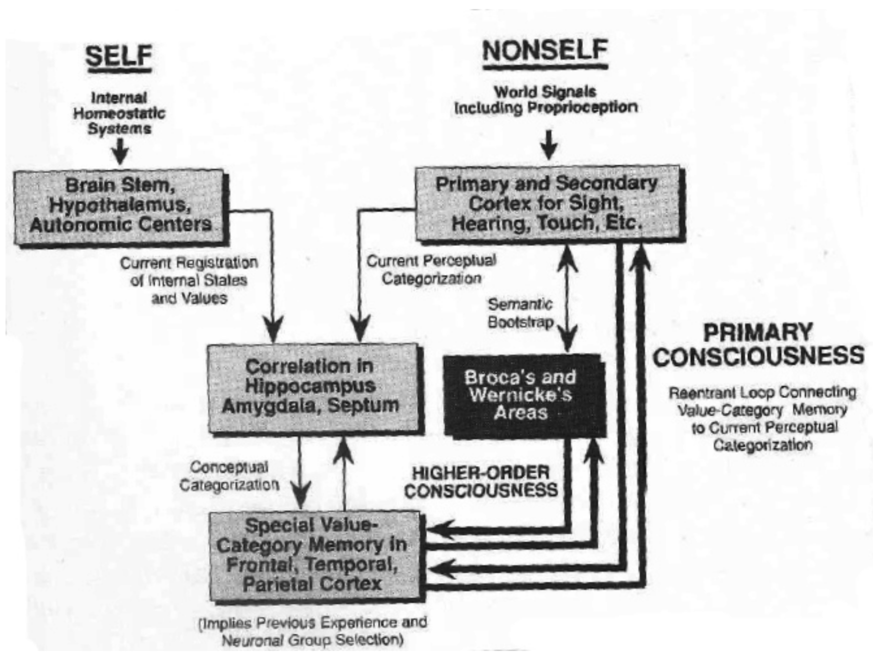


Figure 2A. The distinction between Primary Consciousness and Higher Order Consciousness. The essential difference is the addition of the Broca and Warnicke areas along with their associated reentrant connections (shown in bold). [From Gerald Edelman, *Bright Air, Brilliant Fire: On the Matter of the Mind*].

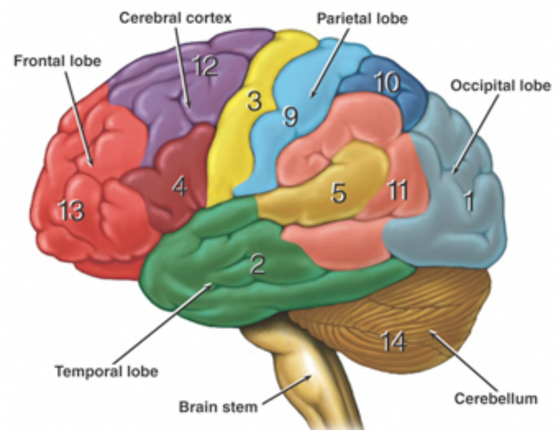


Figure 2B. Some of the regions of the brain referred to in Fig. 2A. Generally, the functions associated with the various numbered areas are: (1) primary visual cortex; (4) Broca's area; (11) Wernicke's area. Broca's area gives us the ability to produce language efficiently and Wernicke's area is associated with the ability to understand spoken language.

Note that the reentrant connections, indicated by the pairs of arrows in opposite directions, are in continuous operation and change to maintain the flow of words, images, etc. For primary consciousness, the lower time limit for the integration of this flow of information is about half a second. The reentrant connection associated with primary consciousness (shown in bold) allows the value-category memory and perceptual categorization to occur in real time. It is this "bootstrapping process" or continuous interaction that is responsible for the rise of primary consciousness.

The regions of the brain upon which primary consciousness depends are evolutionarily old. Edelman estimates that if one assumes that a cortex is required, primary consciousness could be some 300 million years old.

Antonio Damasio, in his chapter titled *A Neurobiology for Consciousness* in *Neural Correlates of Consciousness*—published in 2000 and edited by Thomas Metzinger—defines the concept of Core Consciousness, which is equivalent to Edelman's Primary Consciousness:

> "Core consciousness is a simple, biological phenomenon. It has one level of organization, it is stable across the lifetime of the organism, and it is not dependent on conventional memory, working memory, reasoning, or language. I believe it is not exclusively human. On the other hand, extended consciousness [equivalent to Edelman's Higher Order Consciousness] is a complex biological

> phenomenon. It has several levels of organization, and it evolves across the lifetime of the organism. It depends on conventional memory and working memory, and when it reaches its peak, it depends on language as well. I believe that simple levels of extended consciousness are present in some nonhumans, but extended consciousness attains its maximal development only in humans".

According to Edelman, primary consciousness, which excludes true language ability, is "composed of phenomenal experiences such as mental images, but it is bound to a time around the measurable present, lacks concepts of self, past, and future, and lies beyond direct descriptive individual report from its own standpoint." But there is evidence, that the capabilities of primary consciousness are far more extensive and include many abilities that might be thought to be reserved to higher-order consciousness.

Much of mathematical thought, for example, involves visual imagery and recent research by Giorgio Ganis, William Thompson, and Stephen Kosslyn in their 2004 article in *Cognitive Brain Research* shows that visual imagery and visual perception use the same areas in the brain, although some sensory processes may be used in a different way by visual imagery and perception. Moreover, there is evidence that mathematical reasoning and language is functionally and neuroanatomically independent in adults; the syntax of language (how words are put together to form sentences), whether spoken or written, is distinct from mathematical syntax. Mathematical expressions are not translated into a language format in order to use the specialized syntactic ability for language. Rosemary Varley, et al. put it this way in their 2005 article in the *Proceedings of the National Academy of Sciences*: ". . . language grammar might provide a 'bootstrapping' template to facilitate the use of other hierarchical and generative systems, such as mathematics. However, once these resources are in place, mathematics can be sustained without the grammatical and lexical resources of the language faculty." Elizabeth Brannon, writing about this paper in the same journal and year, said that the work ". . . demonstrates that symbolically mediated representations of numbers can be manipulated in complex ways in human minds that have been robbed of their capacity for linguistic grammar and therein provides additional evidence that mathematics and language are functionally and neuroanatomically independent".

The great mathematician Jacques Hadamard has said that “words are totally absent from my mind when I really think” and, quoting Schopenhauer, maintained that “thoughts die the moment they are embodied by words.” Roger Penrose, another mathematician and one of the most creative people of modern times, has written in his popular book *The Emperor’s New Mind,* “Almost all my mathematical thinking is done visually and in terms of non-verbal concepts, although the thoughts are quite often accompanied by inane and almost useless verbal commentary.” Einstein maintained a similar position on his mode of thought.

Although non-verbal forms of thought may be implicit in primary consciousness and may underlie even verbal thought, language and the ability of humans to communicate may well be necessary for the full development of non-verbal thinking. Scientists, engineers, carpenters, musicians, and many others use a variety of modes of thought depending on the subject matter with which they are dealing. Crows have been known to improvise tools, and myriads of animal owners have seen their pets respond to emergencies in ways that could neither be learned nor instinctive, but rather the result of mental problem solving. The lack of true language—with which higher-level consciousness in humans attains its maximal development—may not be necessary for the simpler forms of higher-level consciousness that appear in many animals. Nonetheless, while the understanding and use of simple syntactical language comprised of up to three word sentences is found in great apes and dolphins, complex syntactical language is only found in humans.

As pointed out by Gerhard Roth in his chapter on *The Evolution and Ontogeny of Consciousness*, contained in the book *Neural Correlates of Consciousness*: The general cytoarchitecture of the human cortex “is indistinguishable from that of other primates and most other mammals. Likewise, no differences have been discovered so far between humans and nonhuman mammals with respect to short-term or long-term plasticity of cortical neurons, the action of neuromodulators, and so on. Only two things have been discovered that could drastically distinguish the human cortex from that of other primates: (1) differences in growth rate and length of period of growth and (2) the presence of the Broca speech center.”

Also involved in the use of speech is the Wernicke area of temporal lobe of the brain that is responsible for understanding the meaning of words and sentences, an ability which also depend on syntax and grammar. A homologous area also exists in the brains of other mammals, so that the relatively recent evolutionary change responsible for allowing humans to develop the capacity to use syntactically complex languages probably occurred in the Broca center.

Higher-Order Consciousness

As can be seen in Edelman's Fig. 2, higher order consciousness arises from the acquisition of true language. How language is learned is represented by the "Semantic Bootstrap" phrase in the diagram. Semantic bootstrapping is intended to correspond to the process where affect, reward, learning, and categorization lead to speech acquisition in the human child. The result of language acquisition in turn allows, to quote Edelman, the "concepts of the self, the past, and the future [to] be connected to primary consciousness. 'Consciousness of consciousness' becomes possible". As we have seen above, this may be too restrictive since elements of higher-order consciousness are found in some animals and mathematical reasoning and syntax are independent of language in adult humans.

Consider first the concept of self. This is usually tested in animals by experiments involving self-recognition in a mirror. Interestingly enough, baboons can recognize group members in a slide without difficulty while monkeys cannot; chimpanzees also demonstrate self-recognition and can also use a mirror to remove marks from their face or body. In human children, self-detection begins at three months while mirror self-recognition begins at eighteen months on average; self-recognition in photos starts at two years and a true "theory of mind" emerges at four years. It is believed that the same sequence of events occurs in apes, although at a much slower rate.

Edelman states in his 2006 book *Second Nature* that: "Dogs and other mammals, if they are aware, have primary consciousness. This is the experience of a unitary scene in a time period of at most seconds that I call the remembered present—a bit like the illumination by a flashlight beam in a dark room. Although they are aware of ongoing events, animals with primary consciousness are not conscious of being conscious and do not have a concept of the past, the future, or a nameable self". Tulving, in his 1983 book *Elements of Episodic Memory*, maintains that "Remembering past events is a universally familiar experience. It is also a uniquely human one. . . . members of the animal kingdom can learn, benefit from experience, acquire the ability to adjust and adapt, to solve problems and make decisions, but they cannot travel back into the past in their own minds".

Recently, however, evidence is accumulating that some animals do have episodic memory and are able to retrieve memories from the past and plan for the future—there is the marvelous example of a zoo-dwelling chimpanzee in Sweden that was found storing stones in the morning to throw at visitors in the afternoon! In their 2001 review article in the *American Journal of Primatology*, Bennet Schwartz and Siân Evans challenge the claim that episodic memory is a uniquely human phenomenon and present evidence that great apes and other primates may indeed possess episodic-like memory. In his 2003 article in *Learning and Motivation*, Thomas Zentall makes the point that "the evidence for subjective time travel in humans is typically based on verbal report and elaboration. Such evidence cannot be obtained from animals. However, we may have indirect evidence for episodic memory and planning. For example, we can show that animals can 'report' about their recent past experience when they are unexpectedly asked to do so—performance that is analogous to episodic memory." And in their 2013 *Proceedings of the National Academy of Sciences* article *The Evolution of Episodic Memory*, Timothy Allen and Norbert Fortin conclude that "Episodic memory is the remarkable capacity to remember specific personal experiences. Although it was originally thought that this capacity was particular to humans, the ample evidence reviewed here indicates that core properties of episodic memory are present across mammals, as well as in birds".

In summary, here is what seems to be the case: The elements of thought of animals including humans are ultimately related to their sense impressions and their neural cognates. They *are* the modes of thought. Episodic memory allows us and some animals to conceive of objects independent of immediate sensory input; and the ability to project thoughts into the past and into the future enables the possibility of imagining objects and possible events as they may have been (beyond the extent of memory) and project the present into the future. While animals may have some capacity for episodic memory and be able to project their thoughts into the past and future to some extent, as far as we know only humans are able to do this with great fidelity. But then there is the question of elephants.

Gerhard Roth put it this way:

> "We have not yet found anything in brain anatomy that would explain the factual or alleged superiority of humans regarding cognition and consciousness. Given the fact that *Homo sapiens* has an absolutely and relatively large brain and cortex, he appears to be the animal with the highest number of cortical neurons and/or synapses, with the probable exception of elephants. . . . They have a similarly enormous brain (around 4 kg) and cortex of about 8,000 cm^2, which is not only thicker than that of cetaceans but also possesses a 'normal' six-layered structure. Assuming that the number of cortical neurons is 2/3 the value found in primates, elephants should have at least as many cortical neurons and cortical synapses as humans. Again, we do not know enough about the organization of the cortex of elephants, but elephants should come close to the cognitive and mental capabilities of man, if only the number of cortical neurons and synapses counted. Perhaps it might be safer to restrict out consideration to the size of the associative cortex, [but] . . . Available data suggest that—contrary to common belief—the associative cortex has increased roughly in proportion to an increase in brain and cortical size".

What then gave humans the enormous advantage they have over other animals? The answer, as so strongly emphasized by Edelman, is true language, but culture also plays an important role. Human culture evolved slowly in several stages. The use and making of tools is perhaps the oldest achievement. Various animals from birds to simians use tools of the simplest sort, such as a stick or rocks, and even pass specific skills on from generation to generation, but not at the level of even the earliest humans. The use of fire also appeared among early humans as well as the use of dogs in hunting—some say the latter represented a type of symbiosis. But most important was the use of language. Some 50,000 to 80,000 years ago there was a flourishing of symbolic use that is believed to be a result of neural changes and reorganization in the brain that resulted in an enhanced ability to use language in a way that greatly improved communication. This created an enormous selective advantage that rapidly spread through the population. The result was the development of sophisticated hunting and gathering cultures that spread throughout the globe. Each showed variations in terms of tools, shelter, and other cultural adaptations to local conditions. Finding food and the activities associated with it left little time for any other activities, but nonetheless these societies developed extensive religions and oral traditions.

The Neural Basis of Human Language

Broca's and Wernicke's areas of the brain were introduced earlier in this essay and their discovery in the late 19th century informed much of the later thinking and modeling of human language. As late as 1970, it was claimed that the neurological basis of human language was formed by the linking of Broca's area with Wernicke's. This model was used by linguists such as Noam Chomsky to argue for an innate linguistic knowledge and to propose the concept of a Universal Grammar. The empirical basis of the model is the observations associated with the aphasias resulting from damage to the Broca and Wernicke areas. Language production is impaired by damage to the motor association cortex that extends to Broca's area. The result is an inability to use proper word order and syntax. Damage to Wernicke's area results in impaired language comprehension.

This led to the rough identification of syntactically complex languages with Broca's area and semantic interpretation with Wernicke's. But then there is the issue of tense. Animals appear to have episodic memory and have some sense of the past and future (remember the chimpanzee with the stones!). This implies that they have at least the foundation for the concept of tense, but because of their very limited language ability it is doubtful that the simple few word combinations they can form include any variation for tense.

Modern research has shown that Broca and Wernecke aphasias involve more than the latter cortical areas. Philip Lieberman in his 2002 book *Human Language and our*

Reptilian Brain; The Subcortical Bases of Speech, Syntax, and Thought has summarized our understanding today as follows:

> ". . . the neural bases of human language are not localized in a specific part of the brain. The brain's dictionary appears to be instantiated by means of a distributed network in which neuroanatomical structures that play a part in the immediate perception of objects and animals as we view them or the gestures associated with tools as we use them are activated. The lexicon appears to connect real-world knowledge with the sound patterns by which we communicate the concepts coded by words. Like other neural structures implicated in language, it is plastic and is shaped by life's experiences. Human beings possess a verbal working-memory system that allows us to comprehend the meaning of a sentence, taking into account the syntactic, semantic information coded in words as well as pragmatic factors. Verbal working memory appears to be instantiated in the human brain by a dynamic distributed network that recruits neural 'computational' resources in response to task demands such as syntactic complexity and sentence length. The neural network that is the basis of verbal working memory links activity in posterior temporal regions of the neocortex, including Wernicke's area, with frontal regions such as Broca's area (Brodmann areas 44 and 45), frontal regions adjacent to Broca's area, the premotor cortex (area 6), the motor cortex, the supplementary motor area, the right-hemisphere homologues of Wernicke's and Broca's area, and the prefrontal cortex. The anterior cingulate cortex, the basal ganglia, and other subcortical structures such as the thalamus and cerebellum also are implicated.
>
> The roots of present human linguistic ability probably go back to our distant ape-human ancestor. Lexical ability and simple syntax probably were present from the start, speech soon afterward. Syntax is not the touchstone of human language. Syntactic ability undoubtedly was present in a limited degree in the earliest hominids. When speech or complex syntax came into being probably will never be known, but some of the necessary neural and anatomical prerequisites were present from the start of hominid evolution."

Homunculus: The "I" or Inner Self

The acquisition of language induces a dichotomy that plays a fundamental role in laying the foundation for acceptance of the mind-body dualism introduced by Descartes. The famous mirror recognition test shows that very young children only begin to recognize themselves in a mirror somewhere around 20-24 months. It is also around this age that the first words begin to be uttered and that a self-concept emerges. Soon they can say "I want . . ." indicating that they understand that what they want is in the external world and the "I" that wants it is different from their perception of the object itself. This is in essence the nature of this second type of dualism that makes a distinction between the perception of an object and the "I" who is the observer; conscious scenes, whether visual, auditory, or based on other perceptual modes, are viewed by a person's inner self.

The duality introduced by Descartes today takes another form and lives on in the extensive philosophical literature in the guise of higher-order theories of consciousness (not to be confused with Edelman's use of the phrase). The idea is that what makes a mental state conscious is that it is the object of some higher-order representation of another mental state. What this actually means leads to many types of higher-order theories of consciousness. An appreciation of this literature can be found in the 2004 book *Higher-Order Theories of Consciousness* edited by Rocco Gennaro.

Gilbert Ryle calls the idea of a higher-order consciousness "Descartes' Myth" and characterizes "The Official Doctrine", which he calls "the dogma of the Ghost in the Machine", simply a "category-mistake". He writes, in his 1949 book *The Concept of Mind*, "My destructive purpose is to show that a family of radical category-mistakes is the source of the double-life theory. The representation of a person as a ghost mysteriously ensconced in a machine derives from this argument. . . . As the human body is a complex organized unit, so the human mind must be another complex organized unit, although one make of a different sort of stuff and with a different sort of structure".

Arthur Koestler, responding to Gilbert Ryle in his 1967 book *The Ghost in the Machine*, called Ryle ". . . a prominent representative of the so-called Oxford School of Philosophy, which, in the words of one of its critics, 'treats genuine thought as a disease' . . . This curious philosophical aberration is now on the wane". Given the continued and ongoing discussion of higher-order theories of consciousness Koestler was clearly correct, but subsequent scientific evidence supports Ryle's position.

Whatever one may feel about this ongoing controversy, what is known today is that the sense of self in the form of a homunculus that views and interprets the ongoing scene when we are conscious is not supported by neurobiology. As put by Damasio, "The sense of self does not correspond, in neurobiological terms or in terms of cognitive operation, to the usual intuitions of self as knower, or of self as homunculus interpreter". Nonetheless, a satisfactory account of consciousness must explain the sense of self that we all have.

Edelman, in his 2003 *Proceedings of National Academy of Sciences* article insists that in animals with primary consciousness the emergent self is not self-conscious and that "Only with the flowering of higher-order consciousness and linguistic capabilities does a self arise that is nameable to itself. Consciousness of consciousness becomes possible via the linguistic tokens that are meaningfully exchanged during speech acts in a community. Episodic memory, which requires the activity of the hippocampus, contributes to the sense of continuity experienced by such a self". Nevertheless, he cautions "the temptation to appeal to a witness, to a homuncular self, must be resisted".

But we have seen that, at least in humans, primary consciousness is capable of very sophisticated non-verbal thought, and higher primates are now known to exhibit the capacity for episodic memory. Consider a visual scene created by primary consciousness. Constructing the scene requires the brain to coordinate and synchronize separate brain regions in an in phase and coherent manner. It has also been shown that cognitive tasks requiring awareness enhance the coherence of the involved brain regions. Given that many sophisticated cognitive tasks including mathematical ones are non-verbal, I would argue that the self that emerges from this process *is* self-conscious, at least in humans, and probably also in the higher mammals.

Try a simple experiment: look out the window without thinking in words or listening to any other form of language. You will be aware of the scene in front of you and also the non-verbal sounds around you as well as many other sense impressions from your body. Are you self-aware? I would argue that you are. This self-consciousness derived from primary consciousness is limited in that there is no sub-vocalization commenting on the scene or performing a verbal analysis; others could call your attention to part of the visual scene by pointing, but communication of the other sense impressions making up the other components of the scene becomes very difficult. So, there is no doubt that the neural changes and reorganization in the brain that resulted in an enhanced ability to use language led to a qualitative change in the nature of primary consciousness and to the sense of self-awareness, but that does not mean that primary consciousness itself was not self-aware in the absence of true language.

The word "scene" is used by Edelman to mean "a *correlation* between different kinds of categorizations"; however, "there is no actual image or sketch in the brain". The "correlation between different kinds of categorizations" is generally known as the "Neural Binding Problem". The representation of sensory input is processed in different regions of the brain and the problem is how the brain binds these distinct neural representations into a single coherent scene.

One hypothesis in the current literature uses the idea that the various oscillation frequencies observed in the brain serve to "synchronize" the various regions of the brain involved in representing the scene. Each oscillation frequency has characteristic spatial range that links specific brain regions to different cognitive functions. The scope of this concept is all inclusive. As put by Leonardo P.G. De Assis in his 2015 article in *Activitas Nervosa Superior* (ANS: *Journal for Neurocognitive Research*) article:

> "According to this classification, alpha waves (8-15 Hz) were associated with all cortical regions, mainly thalamus and hippocampus, with a long-range synchronization. Synchronization in these frequencies is associated with attentional processes. Beta waves (13-30 Hz) were associated with all cortical regions, sub thalamic nucleus, basal ganglia, hippocampus, olfactory bulb, and they also have long-range synchronization. The principal associated cognitive processes are perception, attention, motor control, sensory gating, top-down control, and consciousness. Finally, Gamma waves (30-200 Hz) with their long-range synchrony act in all cortical regions: the hippocampus, retina, tectum, basalganglia, and olfactory bulb . . . . Among the brain functions associated with this frequency are perception, attention, memory, consciousness, synaptic plasticity, and motor control."

De Assis notes that "Although several studies have shown that this proposal has a number of advantages, today no one knows for sure, which specific mechanism the brain uses to accomplish this binding by synchronization".

The binding problem may be a pseudo-problem and simply represent the natural propensity to wish to create what Edelman maintains does not exist: "an actual image or sketch in the brain" that would comprise the "scene" rather than "a *correlation* between different kinds of categorizations". What does this mean? If I look at a red triangle the different areas of the brain that process color and geometrical images are "correlated" and color and shape are different "categorizations" of sensory input. If the triangle is slowly rotated, the shape and color remain the same, but the orientation changes, thereby introducing another area of the brain to be correlated[*] but one that must be updated at a rapid enough rate to give the perception of smooth motion. At any instant of time, a updated scene is perceived and it is these instantaneous correlations themselves that make up the changing scene.

When one is thinking using language, memory serves to recall words, but they are not strung together in a random manner; this mode of thought involves the creation of well-formed sentences that in turn require the use of a grammatical structure, the syntax learned at an early age. The same is true for mathematical or visual thought, but the syntax is different.

A clue to understanding the sense of self that we all have comes from the claim that primary consciousness derives from the experience of a unitary "scene" that lasts for at most seconds, which Edelman calls "the remembered present". This is the ongoing flow and comparison of the "scene" of the *immediately past present* with the *current present*. The "scene" is not just passively formed from the ongoing change in sense perceptions such as visual or oral, but can also include ongoing visualization or sub-vocalization used in thought.[†]

---

[*] Vision is believed to be mediated by three parallel pathways that process information for depth and form, color, and motion.

[†] Sub-vocalization is used when thinking in words and sometimes occurs during reading where it is often thought to limit reading speed.

The *immediately past present* and the *current present* are not experienced as a series of separate static scenes ordered in time (Edelman's flashlight analogy), such as in a slowed down movie film, but as a smooth flowing present, comparable to a movie when the film is moving at a normal speed. This smooth flowing present, which includes thinking using various modes of thought, is the source of our feeling of self; the "homunculus" that is not, it must be emphasized, a higher form of consciousness, particularly in the philosophical sense.

With regard to Edelman's division of consciousness into primary and higher-order, we have seen that primary consciousness, at least in humans, is fully capable of the high-level symbolic thought used in geometric or mathematical thinking; for the language mode of thought, the newer Broca's and older Wernicke areas of the brain are required.

One thing that has not been discussed is motivated behaviors such as curiosity. How, *ab initio*, without external stimulus, do people begin thinking about some topic? It is one thing to respond to a question but quite another to formulate the question when reviewing what one knows about some topic. The ability to do this surely begins in childhood when children, who have an innate motivation to explore, are rewarded for problem solving—such as geometric puzzles—and language acquisition. These activities are quite distinct since, as discussed earlier, it has been found that language syntax and mathematical/geometric syntax are distinct. This difference will inform the examples tentatively introduced here.

Consider a child old enough to have the manipulative skills required who sees someone use an easily handled nut and bolt to hold two objects together. If given a second nut and bolt, children will examine the two pieces and try to manipulate the two so as to screw them together. This behavior can be observed before the child has developed more than rudimentary language skills. When they finally succeed, and if they are rewarded for solving the problem, when they see something similar involving perhaps holding two pieces of wood together, the use of the nut and bolt will come to mind. Similar behavior in an older child will lead to the reward being the pleasure of solving the problem whether mechanical or a mathematical puzzle. The parts of the brain utilized for this type of geometrical/mathematical self-motivated thought are involved more with Edelman's primary rather than higher order consciousness—language and its associated syntax plays little if any role.

In the case of language acquisition, Edelman has described the process as follows: "Phonology provides the means to connect categorized objects to semantics. As reentrant connections are made with concept centers, semantic bootstrapping occurs [see Fig. 2A]. As a lexicon is built and sentences are experienced, the categorization of their arrangements leads to syntax."

Many children a few years old will love the poem

> Twinkle, twinkle, little star,
> How I wonder what you are.
> Up above the world so high,
> Like a diamond in the sky.

They will repeat it endlessly and will often be outraged when confronted with the Mad Hatter's version from *Alice in Wonderland*.

> Twinkle, twinkle, little bat.
> How I wonder where you're at!
> Up above the world you fly,
> Like a tea tray in the sky!

"No, no! That's not right!" The child rejects this not because logically tea trays do not fly but because it is not a faithful rendition of the original. The illogical category error, when understood, only makes the altered version more amusing. Once the rhyme pattern is understood, children may attempt to make up their own versions and get great joy when they succeed. The pleasure is a strong motivation later in life to read and try to write other poetry. Here the syntax of language is involved and this kind of self-motivated thought involves Edelman's higher-order consciousness.

In a child these two modes of thought are generally mixed. When learning to fit a triangle or a square in a puzzle the geometric learning is often coupled with acquiring the words for the geometrical shapes. This allows direct communication with an adult, which can greatly accelerate geometrical learning. Thought precedes language, but is greatly expanded by language acquisition.

Adults may well be conscious of these different modes of thought, as was Roger Penrose when he said. “Almost all my mathematical thinking is done visually and in terms of non-verbal concepts, although the thoughts are quite often accompanied by inane and almost useless verbal commentary.” But such inane verbiage may well link to other useful non-verbal concepts so that one has a synergistic interaction between the two modes of thought as occurs in childhood.

The Ultimate Source of Consciousness

The evolution of self-consciousness from inanimate matter leaves many people uncomfortable. Philosophically, if one accepts the integration of mind and matter the integration of the two requires that one to accept either panpsychism (the doctrine that mind in one form or another is a part of the entire universe including the matter it contains) or that mind is an emergent property arising from inanimate matter.

Almost trivially put, the emergence of mind means the following: The brain is composed of complex networks of neurons that themselves owe their existence to molecules, which in turn are composed of atoms composed of elementary particles and the forces that govern them. If one assumes that molecules are not living and contain no element of “mind”, thus eliminating panpsychism, then neurons are the lowest component of the brain that is living and could be considered to have some element of “consciousness” provided one stretches the meaning of the term. As brains of the animal kingdom become more complex, true consciousness arises as an emergent property. Much of the following discussion of emergence comes from an essay I wrote titled *The Demystification of Emergent Behavior* (http://arxiv.org/pdf/0907.1117).

Emergent behavior that appears at a given level of organization may be characterized as arising from an organizationally lower level in such a way that it transcends a mere increase in the behavioral degree of complexity. It is therefore to be distinguished from systems exhibiting chaotic behavior, for example, which are deterministic but unpredictable because of an exponential dependence on initial conditions. In emergent phenomena, higher-levels of organization are not determined by lower-levels of organization; or, more colloquially, emergent behavior is often said to be "greater than the sum of the parts".

Associated with this phenomenon is a sense of the mysterious: the emergent properties of the collective whole do not in any transparent way derive from the underlying rules governing the interaction of the system's components. Unfortunately, there is not even a universally acknowledged definition of emergence. Nor do the concept and its explication in the literature constitute an organized, rigorous theory. Instead, it is more of a collection of ideas that have in common the notion that complex behavior can arise from the underlying simple rules of interaction.

Ernst Mayr in his monumental 1982 book *The Growth of Biological Thought* characterizes emergence as follows:

> "Systems almost always have the peculiarity that the characteristics of the whole cannot (not even in theory) be deduced from the most complete knowledge of' the components, taken separately or in other partial combinations. This appearance of new characteristics in wholes has been designated as *emergence*. Emergence has often been invoked in attempts to explain such difficult phenomena as life, mind, and consciousness. Actually, emergence is equally characteristic of inorganic systems. As far back as 1868, T. H. Huxley asserted that the peculiar properties of water, its 'aquosity,' could not be deduced from our understanding of the properties of' hydrogen and oxygen. The person, however, who was more responsible than anyone else for the recognition of the importance of emergence was Lloyd Morgan. There is no question, he said, 'that at various grades of organization, material configurations display new and unexpected phenomena and that these include the most striking features of adaptive machinery.' Such emergence is quite universal and, as Popper said, 'We live in a universe of' emergent novelty'. Emergence is a descriptive notion which, particularly in more complex systems, seems to resist analysis. Simply to say, as has been [done][†], that

[†] Original has the misprint "clone".

> emergence is due to complexity is, of course, not an explanation. Perhaps the two most interesting characteristics of new wholes are (1) that they, in turn, can become parts of still higher-level systems, and (2) that wholes can affect properties of components at lower levels. The latter phenomenon is sometimes referred to as 'downward causation'. Emergentism is a thoroughly materialistic philosophy. Those who deny it, like Rensch [Bernhard Rensch, *Biophilosophy* 1971], are forced to adopt pan-psychic or hylozoic theories of matter." [references deleted]

Notice that this concept of "downward causation" is strikingly similar to Edelman's concept of reentry: ". . . reentry in the enormously complex dynamic core distributed to the thalamus and across the cortex was the key integrative event that led to the emergence of conscious experience".

The higher up one goes in a given hierarchy of emergent behavior, the more the organization seems completely independent of the rules determining the behavior of the levels below—which, nevertheless, is not to deny that the higher-order rules are in some sense inherently determined by the properties of the component parts (Mayr's discussion of "Explanatory Reductionism" is relevant here.). But it is the definition of "inherently determined" that contains the essence of the problem.

How can one resolve this conundrum? The answer may lie in the new, internal degrees of freedom that appear as one ascends a hierarchy of emergence. Consider first a simple example from elementary classical mechanics that has relevance to the formation of molecules and hence also to biology. The number of positional degrees of freedom for 2N particles is given by the product of the number of particles and the number of coordinates needed to specify the location of each of the particles in 3-dimensional space. This is $2N \times 3 = 6N$. Now if the particles are combined so as to produce N bonded pairs, with some bonding distance associated with each pair, the number of external degrees of freedom for the pairs is reduced to 3N. However, new degrees of freedom internal to each of the pairs have appeared—the distance between the particles constituting each pair, and the two angles needed to specify the orientation of each pair in 3-dimensional space, a total of three internal degrees of freedom. Notice that the total number of degrees of freedom (3N to locate the pairs in 3-space and 3N "emergent", internal degrees of freedom) has remained constant.

The conservation of the number of degrees of freedom is subtler in quantum mechanics. Take, for example, the case of the helium atom. A helium atom is comprised of a nucleus (considered as a single particle) and two electrons. As separate particles, assumed to be localized in space, the number of degrees of freedom for the three particles is nine. The combination of the three particles to form a helium atom would lead to three degrees of freedom for the location of the nucleus and six internal degrees of freedom consisting of the quantum numbers n, l, and m for each of the electrons. Of course, this is not the whole story since quantum mechanics sets additional constraints on the numerical values of the quantum numbers n, l and m.

Similarly, the structure and variety of all atoms are determined by the rules of quantum mechanics. But the form of the lattice they or their compounds form may depend on additional emergent degrees of freedom such as temperature and pressure reflecting environmental factors. Although one might argue the varieties of structural forms depending on such environmental factors are emergent, it is also possible to argue—at least in principle—that the underlying quantum mechanical rules could take them into account.

An example where this is not possible is the chemistry of saturated hydrocarbons. The rules of quantum mechanics certainly determine the bonding of carbon and hydrogen, and no matter how structurally complex the hydrocarbon, these rules are faithfully obeyed. But the rules of quantum mechanics say nothing about how many carbon atoms may form a chain or whether they form straight chains or branched-chain carbon skeletons. There are emergent degrees of freedom that appear when atoms combine to form these hydrocarbon molecules. It is these emergent degrees of freedom that determine the chemical properties of the saturated hydrocarbons and these chemical properties could well be viewed as an emergent property of a complex system (saturated hydrocarbon molecules) not fully determined by the underlying quantum mechanical rules governing the bonding of hydrogen and carbon atoms.

The emergent rules that govern the chemistry of saturated hydrocarbons are *dependent* on the underlying rules governing the bonding of hydrogen and carbon, but are not *determined* by these rules. That is, the emergent rules cannot be derived from the underlying quantum mechanical rules governing hydrogen-carbon bonding. The difference here is similar to that found in mathematics between *necessity* and *sufficiency*. It is this distinction that should be used to inform the definition of reductionism, particularly in biology with its hierarchical organization, in light of the reality of emergent phenomena.

In the same way, the rules of chemical bonding (again reflecting the rules of quantum mechanics) specify the structure of DNA, but not the sequence of bases. The possible sequences of bases and length of the DNA molecule itself again constitute emergent degrees of freedom not specified by the rules of chemical bonding. Moreover, given that one of four bases is attached to each nucleotide—the basic unit making up the molecule of DNA—the number of possible sequences of bases for n nucleotides is 4^n. This number grows very quickly and each additional sequence could be considered to be a new emergent degree of freedom.

The sequence of bases determines the genes coding for proteins, small RNAs, etc., and it is roughly at this level that the environment begins to play a significant role in the evolution of life through the Darwinian process of variation and selection.[†] But it is not only the set of genes that is responsible for the diversity of animal forms. Of primary importance are differences in gene regulation during ontogeny.

The sequence of bases in DNA contains regulatory code that governs gene expression both in time and location. While this code constitutes another higher-level set of rules,

[†] In terms of the origin of life, molecular evolution and the earliest living creatures were of course also subject to Darwinian variation and selection.

they are rules that have the additional property of being able to change with time in response to environmental selection at the organismal level. The emergent number of degrees of freedom appearing at this level vastly exceeds those at lower hierarchic levels. The whole issue of epigenetics—defined as heritable changes in gene expression not due to changes in base sequence, essentially what allows cells having the same genetic inheritance to make up the variety of cell types comprising an organism—and its role in evolution is still an active area of research.

At an even higher level, while DNA surely determines the structure of living creatures, it would be impossible to derive their social behavior and organization from only the sequence of bases in DNA.

The hierarchy above, starting from elementary particles thus leads to the rules governing ontogeny. Indeed, from an Olympian point of view, life itself may be viewed as an emergent property of matter.

Consciousness and intelligence appear to emerge gradually as the complexity of life increases. Simultaneously, and as a parallel development, a social structure comes into existence. Social behavior can be as simple as that of slime molds when forming a fruiting body, be relatively complex as in the behavior of an ant colony, or be represented by the far more complex behavior of human societies. All appear as forms of emergent behavior.

If the idea that emergent behavior results from the coming into being of new, internal degrees of freedom that arise as one ascends a given hierarchy of emergence is to hold, the inverse should also be true in the sense that a reductionist analysis should eliminate degrees of freedom in the process of descending the hierarchy through reductionist analysis. Here, reductionism is defined as gaining an understanding of a complex system through detailed analysis of the components of the system and their interactions. From the examples of emergence given above, this would seem to be almost trivially true.

In sum, one should view emergence and reductionism as opposite sides of the same coin. Dissecting complex behavior from the top down eliminates internal degrees of freedom in the course of analysis, while emergent phenomena occur when internal degrees of freedom appear when combining component elements into more complex systems. If individual ants are studied to determine their rules of interaction, there is nothing mysterious about the process. But given those rules, one cannot predict the behavior of the colony because the new degrees of freedom that appear in the collective colony cannot be deduced from the rules of interaction—these rules are *necessary* but not *sufficient* to predict the emergent behavior. It is the unexpected consequences of the additional degrees of freedom that appear mysterious.

Emergence in Biological Neural Networks

There has been a great deal of work done to understand and model the behavior of neurons, an example of which is shown in Fig. 3.

Figure 3. The pyloric dilator neuron on the left is filled with a contrast agent and the image was taken with a confocal microscope; it is from unpublished work by D. Bucher. The figure on the right is actually a work of art done by Greg Dunn called *Stomatagastric Ganglion* (Enamel on composition gold and copper). The similarity is striking. More of Dunn's work and a description of his techniques can be found in his article *Etching the Neural Landscape* in the September-October 2014 issue of *American Scientist*.

Neurons respond to a variety of stimuli, but in the laboratory one often uses an electrical stimulus because it is easily controllable. If the current thus introduced into the body of the neuron is above a threshold level, the neuron will "fire" sending a signal down its axon. The information content depends primarily on the frequency of firing and the duration of firing. If one uses electrodes to measure the resulting impulse, called an "action potential", one sees that nerve cells have a variety of properties, as shown in Fig. 4. The lines below the trace of the displayed action potentials correspond to the normal resting potential of about −70 millivolts for the typical neuron. Any level above this line corresponds to an excitatory or depolarizing signal and those below to an inhibitory or hyperpolarizing signal.

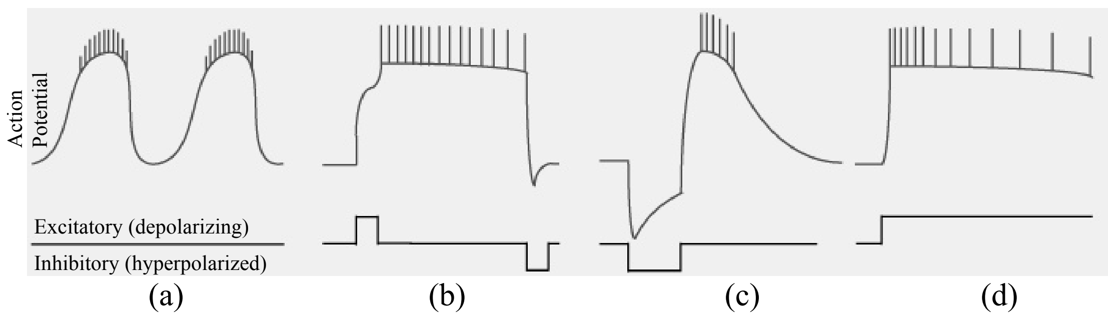
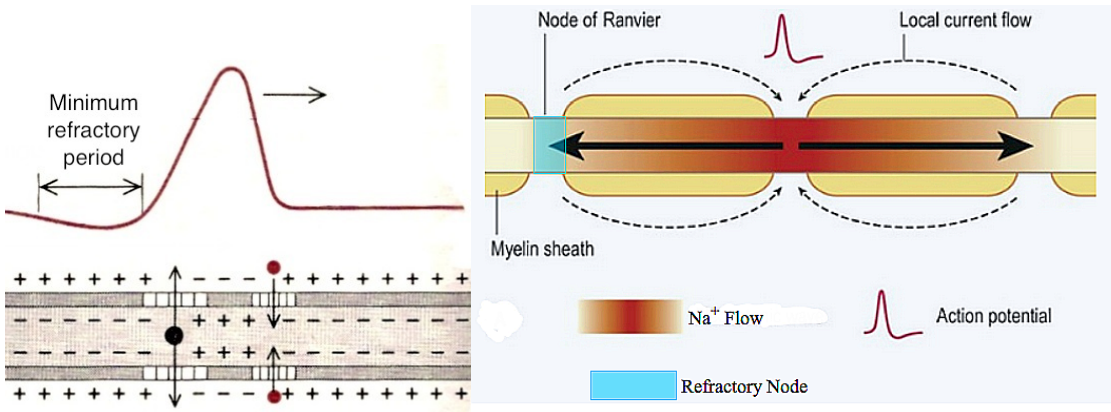


Figure 4. Neuron Properties: (a) endogenous bursting neuron, which has an intrinsically oscillatory membrane potential; (b) plateau potentials that can be terminated by a hyperpolarizing current; (c) rebound firing after inhibation; and (d) spike frequency change during excitation known as frequency adaptation. [Based on Marder and Bucher, *Current Biology* 2001.]

The action potential that travels along the axon of a neuron bears no relation to the current in a conducting wire, but might best be characterized as a propagating depolarization of the charge difference across the membrane of the axon; the interior of the resting cell membrane being negatively charged compared to the exterior environment. Figure 5 shows two types of propagation along the axon.



(a) (b)

Figure 5. (a) shows the mechanism for the propagation of an action potential along and unmyelinated axon. The small filled circles represent Na^+ ions, the larger filled dots K^+ ions; (b) shows the propagation along a myelated axon. [See Fig. 9 for an overall depiction of a neuron.]

The potassium (K) and sodium (Na) ions pass through two types of what are called voltage-gated ion channels. Soon after opening they spontaneously close and will not reopen until the membrane of the axon is repolarized. Action potentials propagate along the axon because a change in voltage in one part of the cell causes the opening of channels in the adjoining part of the cell. The number of ions passing either way through these channels represent only a small fraction of the cell's potassium and sodium ions so that the cell can fire hundreds to thousand of times without having to use energy for the Na^+/K^+ "pump" that generates a high concentration of K^+ ions and a low concentration of Na^+ ions in a resting cell's interior relative to the concentrations in the extracellular medium. It is this "pump" that produces the "resting potential".

Action potentials propagate away from the region where the axon connects to the neuron (the hillock) because of the refractory period of the voltage-gated channels; this refractory period also limits the frequency of the action potentials and thus the information that can be carried by the axon. The depolarization caused by an action potential also results in the opening of the voltage-gated K^+ channels shortly after the depolarization, which restores the resting potential of the axon.

The propagation of the action potential differs in axons that are and are not myelinated. This is seen in Fig. 5 (a) and (b). Myelin sheaths are formed by glial cells and each region of myelin formed by a single glial cell is separated from the next by an unmyelinated region called a node where the axonal membrane is in direct contact with the extracellular medium. This is where one finds the Na^+ voltage-gated channels in this type of axon.

The action potential moves only in the direction away from the neuron because the sodium channels at the previous, passed node are inactivated and thus cannot regenerate another action potential. After a depolarization at one node, the sodium ions diffuse in both directions along the axon causing depolarization at the next activated node to produce an action potential there. In this way the potential propagates along the axon. Note that while there are currents generated by the diffusion of the ions along the myelinated part of the axon between adjacent nodes, the net current along the axon as a whole vanishes. This type of transmission is called "salutatory conduction". Myelination greatly increases the propagation velocity of a signal along the axon, and recent research shows that learning a new motor skill may require active myelination in the brain.[†]

Let us turn to the discussion of neural networks, which will be limited to the behavior of single neurons and the simple pyloric circuit of the crab stomatogastric ganglion that inspired Greg Dunn's artwork shown above in Fig. 3. This will be sufficient for displaying the concept of emergence in neural networks. The reason the crab stomatogastric ganglion was chosen is, to quote Scholarpedia, that:

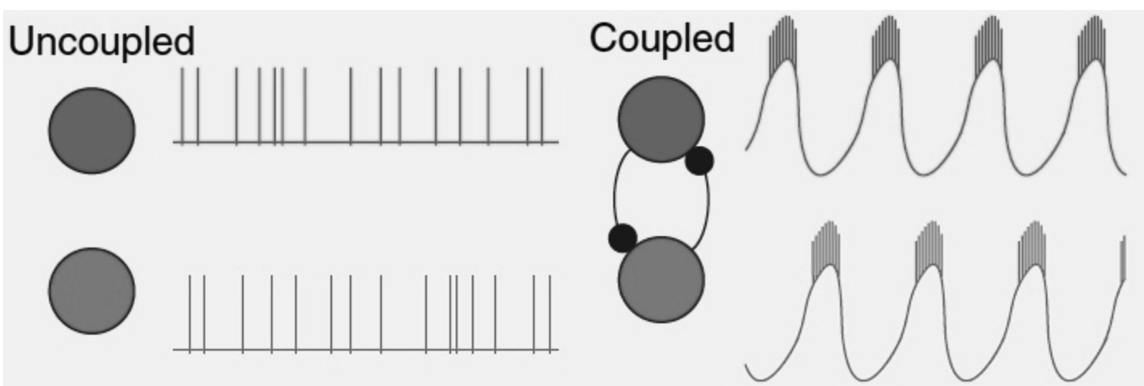
> "The stomatogastric ganglion is a collection of about thirty neurons that sits on the dorsal surface of the foregut (stomach) in decapod crustaceans. The neurons it contains form two central pattern generators (CPGs), namely the pyloric and gastric mill CPGs. The pyloric CPG controls striated muscles that dilate and constrict the pyloric region of the stomach in a cyclic three phase rhythm. The gastric mill CPG produces a slower six phase rhythm that control muscles that produce chewing by three ossicles in the gastric mill. The system has special advantages that make it one of the most well-studied invertebrate neural circuits available. Its importance lies in the fact that the cellular and synaptic properties used by this ganglion are the same as those found in all nervous systems including the mammalian brain but because there are only a few neurons in each circuit the entire 'wiring diagram' for the ganglion has been determined.

[†] A full discussion of the properties of neurons can be found in the 2008 edition of *Molecular Cell Biology* by Harvey Lodish, et al. published by W.H. Freeman and Company.

> Besides, neurons of the stomatogastric ganglion produce cell type-specific voltage output and firing patterns with high consistency among preparations. This feature makes their identification straightforward and helps revealing their connectivity. The stomatogastric ganglion can therefore serve as a model to understand how the synaptic interactions of individual nerve cells produce two functionally different behaviors. It is as a result of these cell-to-cell interactions that sequential spatiotemporal patterns are formed, patterns similar to those produced by CPGs in all nervous systems."

A key reference for understanding the stomatogastric nervous system of lobsters and crabs is the 2007 article in the *Annual Review of Physiology* by Eve Marder and Dirk Bucher titled *Understanding Circuit Dynamics Using the Stomatogastric Nervous System of Lobsters and Crabs*.

Relatively simple emergent behavior can be demonstrated by the coupling of the different types of neurons. The first example is shown in Fig. 6. The large circles represent the neurons while the curves terminating in small dark circles correspond to axons terminating in inhibitory synapses.



(a) (b)

Figure 6. Coupling of two non-rhythmically firing neurons: (a) the firing of each neuron is uncorrelated; (b) the firing of the neurons with reciprocal inhibition. Note that he firing of the two neurons is 180^{o} out of phase due to the reciprocal inhibition. [Based on Marder and Bucher, *Current Biology* 2001.]

By looking at the behavior of the individual neurons in Fig. 6(a), it would seem to be impossible to predict the behavior resulting from synaptically coupling the neurons with reciprocal inhibition. The slow waves generated by reciprocal inhibition coupling are fully synchronized, but the fast action potentials on top of these waves remain unsynchronized. This type of behavior is shown by the pyloric dilator (PD) motor neuron coupled to the lateral pyloric (LP) neuron of the stomatogastric ganglion (STG) of a lobster. The PD neuron is a conditional burster neuron—a burster neuron being one that can respond with an oscillation of the membrane potential so as to produce bursts of action potentials.

The same pattern of firing can be generated by inhibition coupling of an endogenous bursting neuron to a non-rhythmically firing neuron as shown in Fig. 7.

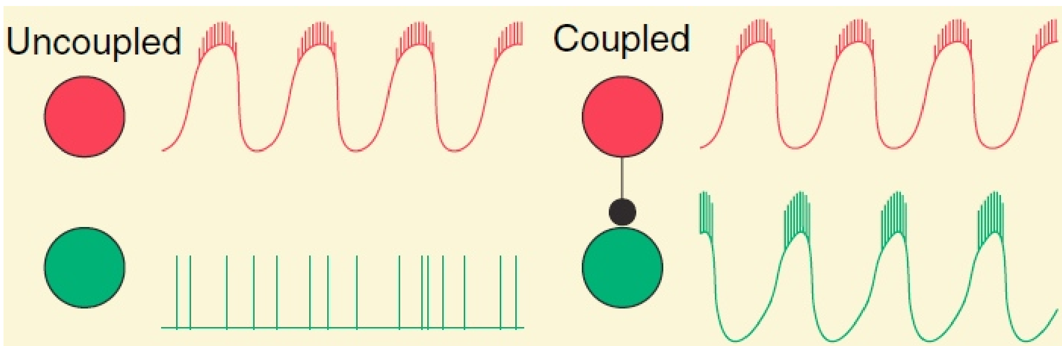


Figure 7. Coupling of an endogenous bursting neuron to a non-rhythmically firing neuron to obtain the same pattern as shown in Fig. 5. [Based on Marder and Bucher, *Current Biology* 2001.]

Again, it would be very difficult to predict the behavior resulting from the synaptic coupling. The final example is the behavior of the STG shown in Fig. 8.

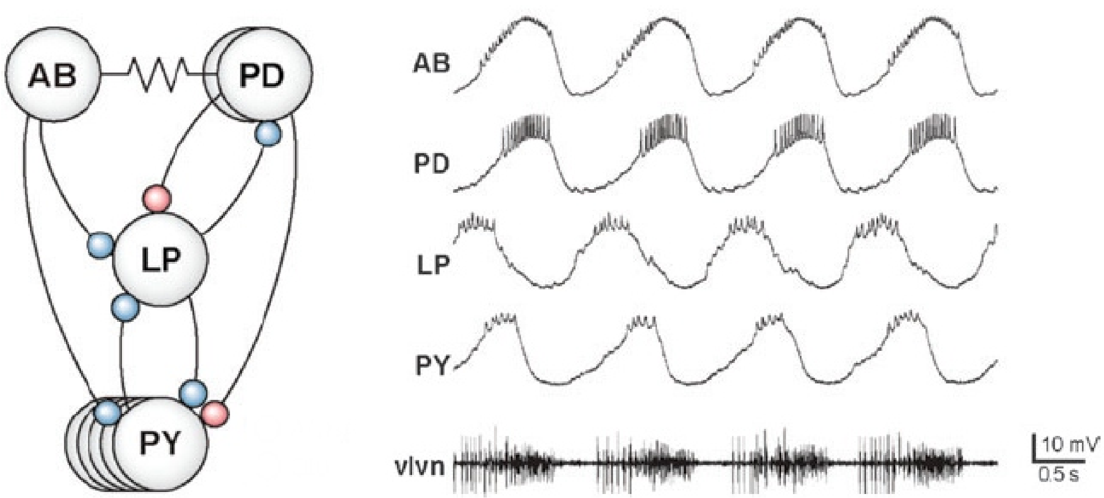


Figure 8. Simplified connectivity diagram of the stomatogastric ganglion (STG): (a) the anterior burster (AB) neuron is an endogenously bursting interneuron that is coupled by an electrical synapse (shown by the resistor symbol) to the PD neurons. LP designates the lateral pyloric neuron and PY the pyloric dilator neuron. All synaptic junctions are inhibiting; (b) simultaneous intracellular action potential recordings from the four types of neurons and an extracellular recording of the ventral branch of the lateral ventricular nerve (vivn). [Based on Marder and Bucher, *Annual Review of Physiology* 2007.]

To get an idea of the number of possible degrees of freedom of a biological neural network one needs at least a notional idea of the actual structure of a neuron. This is shown in Fig. 9. The nerve impulse originates in the cell body and is propagated along the axon, which may have a large number of branches. Such branches are called collateral to distinguish them from the “terminal arborization” ending in presynaptic terminals. Through its collaterally branching an axon may form synapses with up to a thousand other neurons. The number of collateral branches can be used to form the basis for a definition of the degrees of freedom inherent in a biological neural network.

It should be noted that in the context of neural networks the definition of the term “degree of freedom” varies and is used very freely in the literature. There one also speaks of *equivalent degrees of freedom*.

In defining the number of degrees of freedom for DNA, it was seen that given that one of four bases is attached to each nucleotide, the number of possible sequences of bases for n nucleotides is 4^n. In a neural network the question comes down to how many ways can a given neuron attach to other neurons.

Each neuron in a network may be considered to be distinguishable since they will generally have unique connections to their dendrites from other neurons. Therefore, if there are n neurons each of which connects to r others, where $r < n$, the total number of possible ways to pick r neurons out of a total of n is the combination of n things taken r at a time. This is designated by $C(n, r)$.

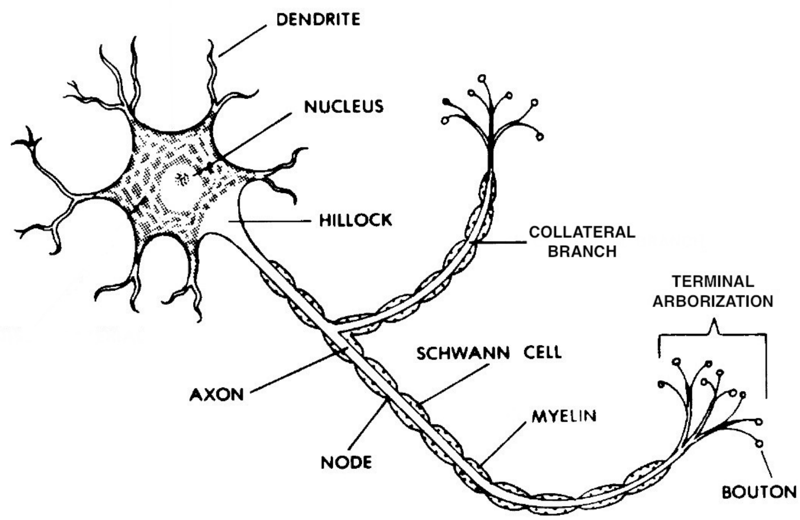
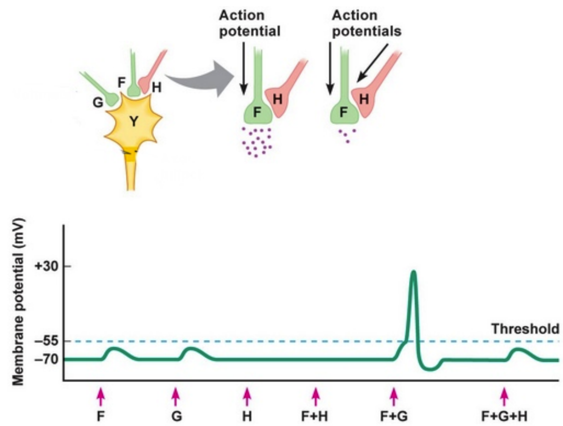


Figure 9. Neurons connect to each other by sending an impulse along axons to the dendrites of the receiving cells, which sums the inputs and if the sum exceeds the threshold potential an action potential is sent down the receiving cell's axon. The axon's collateral branches, which may in turn branch many times along its length, form synapses with as many as 1000 other neurons. Each collateral branch terminates in an arborization (branching) of several to many presynaptic terminals called boutons that attach to the target neuron forming synapses through which impulses pass usually by means of chemicals called neurotransmitters.

For example, if the maximum number of collateral branches is r = 1000, and the total number of neurons that might be reached by these branches is $n = 10^6$, the number of combinations is about 1.5×10^{3432}. This is for each neuron! And in reality the situation is complicated by other factors; for example, there is also the possibility of presynaptic inhibition and facilitation. This is where an inhibitory or facilitory neuron forms a synapse on the terminal of the presynaptic neuron. This is shown for presynaptic inhibition in Fig. 10.



Presynaptic Inhibition

Figure 10. Presynaptic Inhibition. In Presynaptic Inhibition, the membrane potential of neuron Y will exceed the threshold value if both axons F and G carry an action potential from a presynaptic neuron (indicated by the arrows) and axon H does not.

In computing C(n, r) I used $n = 10^6$ and $r = 1000$; is this reasonable? Well, the neural density in the primate cortex has been estimated to be about 100,000 per cubic millimeter. So $n = 10^6$ neurons corresponds to 10 cubic mm or a cube a little over 2 mm on a side, not unusual for the volume occupied by collateral axon arborization. The number of degrees of freedom that become available as the number of neurons in a neural network increases is seen to be simply enormous.

Directly connecting these degrees of freedom to the emergent properties of neural networks is essentially impossible as shown by the previous simple examples of coupled neurons.

Artificial Intelligence

If life and the evolution of intelligence are to be considered to be emergent properties of matter, one is faced with the question of whether or not it is possible to create an "artificial intelligence", and AI (as it is called), has developed into a very large somewhat incoherent field of endeavor. The name itself leaves one wondering why if a true intelligence is created it is artificial; presumably, what is meant by "artificial" is not being based on living, biological components. Often the discussion revolves around digital computers based on the concept of a Turing machine—named after Alan Turing, which is a simplified, hypothetical device that can simulate the logic of digital computer algorithms.

From the discussion in the last section is should be clear that biological neural networks and brains do not operate like digital computers. Nonetheless, digital computers are fully capable of modeling analog circuits in general and neurons in particular. To model a biological neural network it is necessary to fully understand and define in electrical and chemical terms the functioning of the various types of neurons. It is also necessary to model synapses and their ability to form both short and long-term memory.

The synaptic basis for long-term memory is very different from that of short-term memory. Short-term memory strengthens existing synapses; in long term memory, arising from repeated stimulation, a neuron can grow additional synapses enhancing synaptic transmission. Thus, the number of synapses on neurons, and in the brain as a whole, is neither fixed nor static.

Whereas Fig. 10 shows presynaptic inhibition, Fig. 11 shows the mechanism for of presynaptic facilitation and how repeated stimulation causes the nucleus of the cell to express specific genes to synthesize the new proteins needed for synapse growth. It is this formation of new synapses that underlies long-term memory.

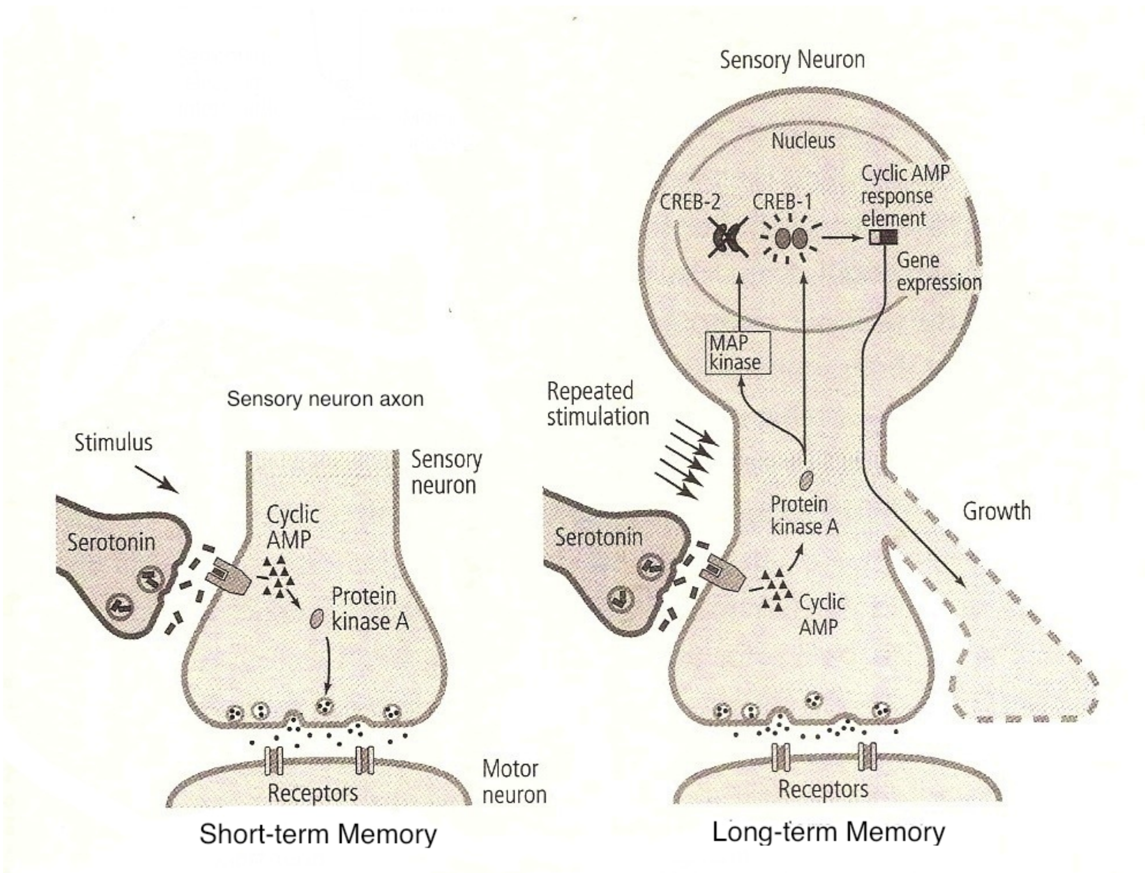


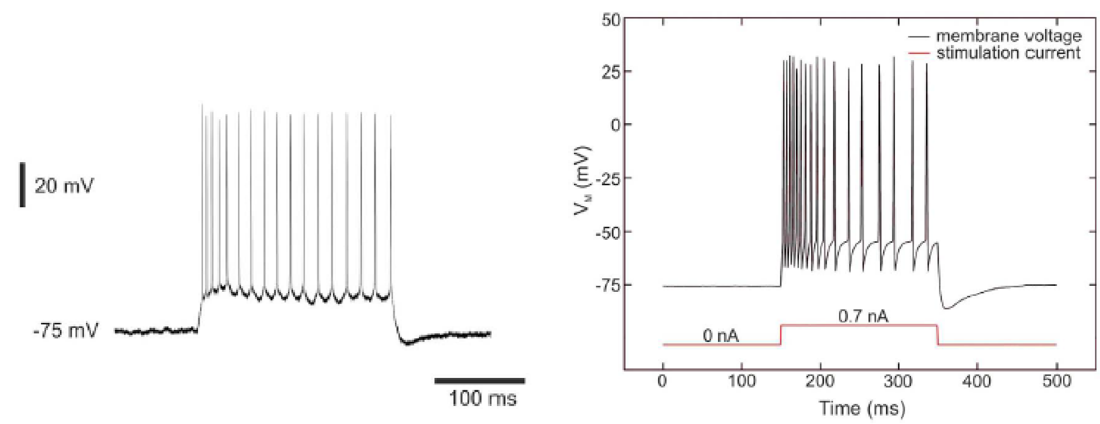
Figure 11. Molecular mechanisms associated with the synapse for short and long-term memory. Serotonin is a neuro-transmitter; cyclic AMP is a cyclic form of adonosine monophosphate that plays a role in controlling enzyme-catalyzed processes in cells; CREB stands for cyclic AMP response element binding protein; and kinases modify proteins by adding a phosphate molecule, a process called phosphorylation. Repeated stimulation tells the nucleus of the neuron to synthesize proteins leading to the growth of new synaptic connections thus increasing synaptic transmission. [Adapted from Fig. 19.1 of Eric Kandel's 2006 book *In Search of Memory*.]

In principle then, it may ultimately be possible to simulate a biological brain with a digital computer, but—although very useful for heuristic purposes—this is a very clumsy approach for creating a non-biological intelligence. Far more promising is something along the lines of the DARPA SyNAPSE program.[†]

SyNAPSE stands for Systems of Neuromorphic Adaptive Plastic Scalable Electronics. The point of the program is to mimic the neurological architectures of biological neural networks and ultimately of brains. The whole endeavor has become known as

[†] http://www.artificialbrains.com/darpa-synapse-program - neurosynaptic-core.

neuromorphic engineering. To do this one must be able to simulate with electronics the behavior of a neuron. An example of this type of work has been given by Filippo Grassia, et al. in their 2011 article in *Frontiers in Neuroscience* titled “Tunable neuromimetic integrated system for emulating cortical neuron models”. Their paper deals with implementing prototypical neurons of the neocortex in very large-scale integrated (VLSI) electronic microchips. Figure 12, adapted from this paper, shows a comparison between a biological regular spiking neuron and its electronic simulation.



(a) (b)

Figure 12. (a) shows the intracellular recording of a regularly spiking neuron from a ferret visual cortex and (b) the response of its electronic simulation. The biological neuron’s response was from a 0.7 nA depolarizing current, the same as that for the simulated neuron. Note that the firing frequencies at the beginning of the activity are not identical.

The brains of living creatures are not an arbitrarily connected neural network but consists of a highly organized structures determined by genetics with many homological components across species. As mentioned earlier, DNA does not have the information carrying capability to specify the enormous number of interconnections of the brain, and the microstructure of the brain that result in functional neural networks is a consequence of learning and Neural Darwinism.[†] A review of “cognitive computing” has been given by Dharmendr Modha, et al. in the August 2011 issue of *Communications of the ACM*. In their article they maintain that “Cognitive computing will lead to novel learning systems, non-von Neumann computing architectures, programming paradigms, and applications that integrate, analyze, and act on vast amounts of data from many sources at

[†] For this history of this concept see: http://faculty.ed.uiuc.edu/g-cziko/wm/05.html.

once”, thus mimicking brain structure. Their work was done under the aegis of the SyNAPSE program.

There has also been some work in the use of living biological neural networks to control mechanical systems, one of the most well covered in the media being the use of rat cortical neurons to control a flight simulator. This was done by Thomas DeMarse and Karl Dockendorf, who describe their work as follows: “A system was created in which a network of living rat cortical neurons were slowly adapted to control an aircraft’s flight trajectory. This was accomplished by using high frequency stimulation pulses delivered to two independent channels, one for pitch, and one for roll. This relatively simple system was able to control the pitch and roll of a simulated aircraft”. [†]

Coda and Implications

I have attempted in this essay to discuss the nature of consciousness and provide an introduction to its material basis. From what is now known about the subject, one is ineluctably driven to accept the realization that consciousness is a natural emergent biological phenomena founded on the the ability of matter to form the molecules needed for life to exist.

This does not diminish humanity or its place in the universe. On the contrary, as put by Homer Smith, quoted in the beginning of this work, “as an intelligent animal, [man] has reason to be proud because he is the first who can ask himself, ‘Whither, Why, and Whence?’ and confident because he can know himself as a creature of the earth who has risen by his own efforts from a low estate”.

The implications of the natural biological basis of consciousness for religious thought are enormous. While one may retain the concept of the soul, provided it is defined as the unique form of consciousness that develops in each individual over their lifetime, one must give up the concept that a soul is immortal, and with it the parallel universe of ghosts, angels, and devils. One might think of these religious constructs and their role in

---

[†] http://neural.bme.ufl.edu/page12/page1/assets/NeuroFlght2.pdf

human life as an allegory for the nature of humanity where each person is a mix of good and evil however these terms may be defined. The origin of these concepts derives from the revelatory form of western religions. Non-western religions may have analogous beliefs that significantly differ from those of the west.

To understand how scientific knowledge conflicts with beliefs based on revelation, it is necessary to understand exactly what is meant by this term. Judaism, Christianity, and it is not unreasonable to include Islam here, are religions of revelation. By revelation, most people are referring to theophany, the sudden and dramatic manifestation or appearance of God or the unveiling of a mystery. Moses seeing the burning bush, or coming down from the mountaintop with the Ten Commandments are examples. Religious fundamentalists, whether Islamic, Christian, or Jewish, believe in theophany. And theophany, as recorded in scripture, gives an absolute, eternal form of truth. Nothing that comes after can alter such truths. It is for this reason that fundamentalists also believe in a literal interpretation of scripture.

As theophany, revelation is the exact antithesis of scientific knowledge; there are no absolute or eternal truths in science. Secular, scientific reasoning cannot accept the divine, with its immutable truths if it is to remain true to itself.

There is, however, a form of revelation—not based on theophany—that *is* compatible with science. As put by James Carroll—a former Catholic priest—in his brilliant history, *Constantine's Sword*, "the truth of our beliefs is revealed in history, within the contours of the mundane, and not through cosmic interruptions in the flow of time. Revelation comes to us gradually, according to the methods of human knowing. And so revelation comes to us ambiguously. Certitude and clarity are achieved only in hindsight, and even then provisionally." Since it is this provisional nature of knowledge that is also the essence of scientific knowledge, religious people who find themselves able to accept Carroll's definition of revelation should have no difficulty accepting the findings of modern science—those findings reflect the will of God.

Accepting that the immortal soul does not exist does not mandate that the religious or "spiritual" dimension of human existence must also be lost. Nor need we lose the religious heritage formed over millennia, or the beautiful art and music engendered during the course of this religious evolution. Science and religion only conflict insofar as religious precepts are at odds with well-established science taken in its broadest sense. Much of the perceived incompatibility between religious thought and science is due to a confusion over their ends and methodology: science only addresses the question of "How?"; it does not, and cannot, answer the question "Why?". It does, however, allow us

> "To see a World in a Grain of Sand
> And a Heaven in a Wild Flower,
> Hold Infinity in the palm of your hand
> And Eternity in an hour."
>
> —William Blake *Auguries of Innocence*

The universe is now known to have had a beginning some 14 billion years ago. Scientists usually use the term (originally introduced derogatorily by Fred Hoyle) "the big bang" for this event; for those with a religious orientation, it is know as the "creation". Despite understanding the evolution of the universe from very shortly after it came into existence, the actual event of its doing so is not understood. This is primarily due to the constraint that its coming into existence is *sui generis*, and the scientific method that relies on the process of the interplay of theory and experiment does not apply. This does not, however, impede scientific speculation.