

## CHAPTER 23

# Machines and Consciousness

*Subhash Kak*

### 1. INTRODUCTION

The mainstream view of science is based on the paradigm of the machine, in which the larger mechanism is built out of smaller subsystems. According to this paradigm, systems can be separated and any process is reducible to a collection of more fundamental sub-processes and, in principle, the machine's future response is completely describable if its current state and inputs are known. Wholes are constructed out of parts, although it is acknowledged that as objects come together in complex assemblies, new properties emerge. Thus chemical properties emerge as elementary particles are brought together in large numbers, and chemical substances in complex systems lead to biological systems. Given that consciousness is a characteristic of higher life, it follows that it is an emergent property.

From the point of view of machine design, one can imagine that ordinary machine behaviour may lead to consciousness, after machine complexity has crossed a critical threshold. But since machines only follow instructions, it is not credible that they should suddenly, just on account of a greater number of connections between certain computing units, become endowed with self-awareness. To speak of consciousness in the machine paradigm is a contradiction in terms. If a machine could make true choices (that is not governed by a random picking between different alternatives), then it has transcended the paradigm because its behaviour cannot be described by any mathematical function.

If one accepts that machines will never become self-aware, one may ask why the brain-machine is conscious, whereas the silicon-computer is not. On a process level, one might ascribe this difference to the fact that the brain is a self-organizing system which responds to the nature and quality of its interaction with the environment, whereas computers don't do that. But other ecological systems, which are biological communities that have complex interrelationships amongst their components, are self-organizing, without being self-aware. This suggests that while self-organization is a necessary prerequisite for consciousness, it is not sufficient.

Yet another possibility is that current science, even when it considers self-organization and special structures of the brain, does not capture the essence of consciousness. Our

scientific framework may be incomplete in a variety of ways. For example, we may not yet have discovered all the laws of nature, and our current theories may need major revision that has implications for our understanding of consciousness.

In truth, objective knowledge consists of many paradoxes. Accumulation of knowledge often amounts to making ad hoc choices in the underlying formal framework to conform to experience. The most fundamental, and a very ancient, antinomy is that between determinism and free will. Formal knowledge can at best be compared to a patchwork. The puzzle is: How is a part of the physical world which generates the mental picture which, in turn, creates the scientific theory able to describe nature so well? Why is mathematics so unreasonably effective?

Cognitive scientists and biologists have considered evolutionary aspects related to cognitive capacity, where consciousness is viewed as emerging out of language. Linguistic research on chimpanzees and bonobos has revealed that although they can be taught basic vocabulary of several hundred words, this linguistic ability does not extend to syntax. By contrast, small children acquire much larger vocabularies—and use the words far more creatively—with no overt training, suggesting that language is an innate capacity.

According to the nativist view, language ability is rooted in the biology of the brain, and our ability to use grammar and syntax is an instinct, dependent on specific modules of the brain. Therefore, we learn language as a consequence of a unique biological adaptation, and not because it is an emergent response to the problem of communication confronted by ourselves and our ancestors. Deaf children provide us a context in which the creative development of a language may be seen. When such children are not taught a sign language, they spontaneously create their personal signs, slowly adding grammatical rules, complete with inflection, case marking, and other forms of syntax.

It has been suggested that human language capacities arose out of biological natural selection because they fulfil two clear criteria: an extremely complex and rich design and the absence of alternative processes capable of explaining such complexity. Other theories look at music and language arising out of sexual selection. But, howsoever imaginative and suggestive these models might be, they do not address the question of how the capacity to visualize models of the world that are essential to language and consciousness first arise. There is also the difficulty that certain cognitive capacities appear in wholes, quite like the Sāṃkhyan *tattvas*.

Finally, there is the Vedic critique of the enterprise of seeking a theory of consciousness.<sup>1</sup> According to this critique, all that *normal* science (*aparā vidyā*) can hope to achieve is a description of objects. But consciousness (*parā vidyā*) is a property of the subject, the experiencing 'I', which, owing to its nature, forever lies outside the pale of *normal* science. The experimenter cannot turn his gaze upon himself, and ordinary reality must have a dual aspect. This duality means that the world of objective causality is incomplete, creating a fundamental paradox: If objects are described by normal science, why is that science not rich enough to describe the body associated with the experiencing subject? The Vedic view recognizes this problem as basic to the experience of reality.

In this article, we mainly consider evidence that negates the view that the brain is an *ordinary* machine. We argue that even with self-organization and hitherto unknown quantum characteristics one cannot explain the capacities associated with the brain. We examine the limitations of computational models by considering questions raised by new

researches in animal intelligence. We also examine the central role of recursive self-organization, as seen, for example, in superorganisms.

## 2. BRAIN AND MIND

The question of consciousness is connected to the relationship between brain and mind. Reductionism takes it that they are identical, and mind is only the sum total of the activity in the brain, viewed at a suitable higher level of representation. Opposed to this is the viewpoint that although mind requires a physical structure, it ends up transcending that structure.

The mind processes signals coming into the brain to obtain its understandings in the domains of seeing, hearing, touching, and tasting using its store of memories. But a cognitive act is an active process where the selectivity of the sensors and the accompanying processing in the brain is organized based on the expectation of the cognitive task and on effort, will and intention. Intelligence is a result of the workings of numerous active cognitive agents.

The reductionist approach to artificial intelligence (AI) emerged out of an attempt to mechanize logic in the 1930s. In turn, AI and computer science influenced research in psychology and neuroscience and the view developed that a cognitive act is a logical computation. This appeared reasonable as long as classical computing was the only model of effective computation. But with the advent of quantum computing theory, we know that the mechanistic model of computing does not capture all the power of natural computation.

Schrödinger spoke of the *arithmetic paradox* related to the mind as being ‘the *many* conscious egos from whose mental experiences the *one* world is concocted’. He adds that there are only two ways out of the number paradox. ‘One way out is the multiplication of the world in Leibniz’s fearful doctrine of monads: every monad to be a world by itself, no communication between them; the monad “has no windows”, it is “incommunicado”. That none the less they all agree with each other is called “pre-established harmony”. I think there are few to whom this suggestion appeals, nay who would consider it a mitigation at all of the numerical antinomy. There is obviously only one alternative, namely the unification of minds or consciousnesses. Their multiplicity is only apparent, in truth there is only one mind. This is the doctrine of the Upaniṣads.’<sup>2</sup>

In the view that consciousness is one of the grounds of reality, together with space, time and matter, consciousness and space-time-matter are complementary because consciousness needs the support of matter and without observers it is meaningless to speak of a universe. The idea of the physical world is constructed out of mental experiences. If I give primacy to this mental experience then I am an idealist, but if I give primacy to the contents of this mental experience then I am a materialist. If I believe that both these have an independent existence then I am a dualist.

If we take it that we know the basic laws of nature and also accept that classical machines cannot be conscious, one must assume that quantum processing in the brain, given appropriate brain structures, leads to awareness. Different states of consciousness such as wakefulness, sleep, dream-sleep, coma have distinct neurochemical signatures, and these different states may be taken to be modifications caused by the neural circuitry on a basic state of consciousness.

The case that quantum theory is the explanation for the power of animal intelligence and basic to biological information processing has the following elements:<sup>3</sup>

1. *Philosophical.* At the deepest level of description nature is quantum-mechanical. The world of mathematics, as a product of the human mind, sits on top of the sequence physical → chemical → mental → ideational (mathematical).  
Since our ideas (dressed in a mathematical form) are able to describe the quantum-mechanical physical reality, the power of information processing in the brain should equal the power of quantum mechanics. Another argument is that quantum mechanics as a universal theory should apply to information and organization and therefore the information processing of the brain cannot be understood but in quantum-mechanical terms.
2. *Neurophysiological.* The interior of living cells is organized around the cytoskeleton which is a web of protein polymers. The major components of the cytoskeleton are the microtubules, that are hollow tubes 25 nm in diameter, consisting of 13 columns of tubulin dimers arranged in a skewed hexagonal lattice. Some researchers have argued that the microtubules support coherent, macroscopic quantum states. They see brain processing as a hybrid quantum/classical computation.
3. *Self-awareness.* Awareness implies conscious choice and this has been compared with a reduction to one-out-of-many possibilities of quantum mechanics. More directly, since there is no credible reason that awareness is a result of the degree of complexity of neural mechanisms doing classical computing, it is reasonable to take it as a fundamental attribute of reality which is manifested by neural hardware running a quantum process. The notion of 'self', which provides a unity to experience, is a consequence of favourable neural hardware tapping the ground consciousness.
4. *Behavioural science.* Human and nonhuman animal intelligence appear to have features that lie beyond the capacity of the most powerful machines. Conceptualization is not unique to humans and ability to use language is not a precondition to cognition or abstract processing. Since we associate linguistic analysis with classical logic, one may presume that cognition is based upon some non-computable program. Intelligent behaviour may be viewed as adaptation to a changing environment. Paralleling adaptive behaviour is the continual self-organization in the brain. Analogously, a quantum system may be viewed as responding to its measuring apparatus (environment) by changing its state. Although non-quantum models for self-organization may be devised, only a quantum approach appears to satisfy different attributes of mental activity.

Many ancient cultures recognized the limitations of mechanistic logic in understanding the autonomy of individuals. In India, the Vedas declare reality as transcending the subject-object distinction and then self-consciously speak of their own narrative as dealing with the problem of consciousness. Specifically, the texts speak of the cognitive centres as individual, whole entities which are, nevertheless, a part of a greater unity. The cognitive centres are called the *devas* (gods), or points of light.<sup>4</sup> The *devas* are visualized in a complex, hierarchical scheme, in which some of them are closer to the autonomous processes of the body and others are nearer creative centres. Mirroring the topology of the outer reality, the inner space of consciousness is seen to have a structure. The Vedic

texts divide the capacities of the mind in various dichotomies, such as high and low, left and right, and so on.

Parallels between the Vedic view and quantum theory are well known. For example, both suggest that reality is consistent only in its primordial, implicate form. The Vedas insist that speech and sense-associations cannot describe this reality completely. It is not well known that Vedic ideas had an important influence on the development of quantum mechanics as described at length in Moore's biography<sup>5</sup> of Erwin Schrödinger, who was one of the creators of the subject.

Vedic narrative insists that ordinary linguistic descriptions of reality will be paradoxical. In quantum physics also use of ordinary logic or language leads to paradoxes such as: the present can influence the past, effects can travel instantaneously, and so on.

The Vedic model of mind provides a hierarchical structure with a twist that allows it to transcend the categories of separation and wholeness. In it, the lowest level is the physical world or body, with higher levels of interface forces, the mind, intuition, emotion, with the *universal self* sitting atop. The lower levels are machine-like whereas the self is free and conscious. The individual's idea of consciousness arises out of associations with events, illuminated by the consciousness principle.

The most striking part of this model is the nature of the *universal self*. Considered to transcend time, space and matter, the *self* engenders these categories on the physical world. Mind itself is seen as a complex structure. Although it is emergent and based on the physical capacities of the brain structures, it cannot exist without the *universal self*. One implication of these ideas is that machines, which are based on logic, can never be conscious.

The Vedic theory of mind is part of a recursive approach to knowledge. The Vedas speak of three worlds, namely the physical, the mental, and that of knowledge. Consciousness is the fourth, transcending world. There is also reference to four kinds of language: gross sound, mental imagery, gestalts, and a fourth that transcends the first three and is associated with the *self*.

These questions are examined in the later Vedic tradition within the frameworks of both Vaiṣṇhnavism and Śaivism. Of the latter tradition, the later Kashmir Śaivism of Vasugupta (AD 800) has in recent years received considerable attention. The 25 categories of Sāṃkhya form the substratum of the classification in Kashmir Śaivism. The Sāṃkhya categories are:

- (i) *five elements of materiality*, represented by earth, water, fire, air, ether;
- (ii) *five subtle elements*, represented by smell, taste, form, touch, sound;
- (iii) *five organs of action*, represented by reproduction, excretion, locomotion, grasping, speech;
- (iv) *five organs of cognition*, related to smell, taste, vision, touch, hearing;
- (v) *three internal organs*, mind, ego, and intellect;
- (vi) and inherent *nature (prakṛiti)*, and *consciousness (puruṣa)*.

These categories define the structure of the physical world and of agents and their minds. Kashmir Śaivism enumerates further characteristics of consciousness:<sup>6</sup>

- (vii) *sheaths or limitations of consciousness*, being time (*kāla*), space (*niyati*), selectivity (*rāga*), awareness (*vidyā*), creativity (*kalā*), self-forgetting (*māyā*), and
- (viii) *five principles of the universal experience*, which are correlation in the universal experience (*sadvidyā, śuddhavidyā*), identification of the universal (*īśvara*), the

principle of being (*sādākhya*), the principle of negation and potentialization (*śakti*), and pure awareness by itself (*śiva*).

The first 25 categories relate to an everyday classification of reality. The next 11 categories characterize different aspects of consciousness that are to be understood in a sense different to that of mental capacities (categories 21, 22, 23). One of these mental capacities is akin to artificial intelligence, which is geared to finding patterns and deciding between hypotheses, but categories 26 through 36 deal with interrelationships in space and time between these patterns and deeper levels of comprehension and awareness.

Various Indian philosophical schools describe the Vedic theories of mind in different expositions. In Vaiśeṣika, the mind is considered to be atomic and of point-like character, anticipating Leibniz's theory of monads, although it is different from it in a crucial sense. According to the Vedic traditions, mind itself must be seen as a complex structure. Whereas mind is emergent and based on the capabilities of neural hardware, it cannot exist without the universal self. One implication of these ideas is that machines, which are based on classical logic, can never be conscious.

It is not well known that the Vedic model had an important influence on the development of quantum mechanics. In 1925, before his creation of wave mechanics, Erwin Schrödinger wrote:<sup>7</sup> 'This life of yours which you are living is not merely a piece of this entire existence, but in a certain sense the "whole"; only this whole is not so constituted that it can be surveyed in one single glance. This, as we know, is what the Brahmins express in that sacred, mystic formula which is yet really so simple and so clear: (*tat tvam asi,*) this is you. Or, again, in such words as "*I am in the east and the west, I am above and below, I am this entire world.*"'

Schrödinger's influential *What is Life?* was also informed by Vedic ideas.<sup>8</sup> This book inspired Francis Crick to search for the molecule that carries genetic information, which led ultimately to the discovery of the structure of the DNA. Schrödinger also believed that Sāṃkhyan *tattvas* could be the only explanation for the developed of sensory organs in animals. According to his biographer Walter Moore, there is a clear continuity between Schrödinger's understanding of Vedānta and his research:<sup>9</sup> 'The unity and continuity of Vedānta are reflected in the unity and continuity of wave mechanics. In 1925, the world view of physics was a model of a great machine composed of separable interacting material particles. During the next few years, Schrödinger and Heisenberg and their followers created a universe based on superimposed inseparable waves of probability amplitudes. This new view would be entirely consistent with the Vedāntic concept of *All in One*.'

### 3. LANGUAGES OF DESCRIPTION

Progress in science is reflected in a corresponding development of language. The vistas opened up by the microscope, the telescope, tomography and other sensing devices have resulted in the naming of new entities and processes.

The language for the description of the mind in scientific discourse has not kept pace with the developments in the physical sciences. The mainstream discussion has moved from the earlier dualistic models of common belief to one based on the emergence of mind from the complexity of the parallel computer-like brain processes. The two old paradigms of determinism and autonomy, expressed sometimes in terms of separation

and interconnectedness, show up in various guises. Which of the two of these is in favour depends on the field of research and the prevailing fashions. Although quantum theory has provided the foundation for physical sciences for 70 years, it is only recently that holistic, quantum-like operations in the brain have been considered. This fresh look has been prompted by the setbacks suffered by the various AI projects and also by new analysis and experimental findings.

The languages used to describe the workings of the brain have been modelled after the dominant scientific paradigm of the age. The rise of mechanistic science saw the conceptualization of the mind as a machine. Although the neural network approach has had considerable success in modelling many counterintuitive illusions, there exist other processes in human and nonhuman cognition that appear to fall outside the scope of such models. Briefly, the classical neural network model does not provide a resolution to the question of binding of patterns: How do the neuron firings in the brain come to have specific meanings or lead to specific images?

Considering that the physical world is described at its most basic level by quantum mechanics, how can classical computational basis underlie the description of the structure (mind) that in turn is able to comprehend the universe? How can machines, based on classical logic, mimic biological computing? One may argue that ultimately the foundation on which the circuitry of classical computers is based is at its deepest level described by quantum mechanics. Nevertheless, actual computations are governed by a binary logic which is very different from the tangled computations of quantum mechanics. And since the applicability of quantum mechanics is not constrained, in principle, by size or scale, classical computers do appear to be limited.

Why cannot a classical computer reorganize itself in response to inputs? If it did, it would soon reach an organizational state associated with some energy minimum and would then stop responding to the environment. Once this state has been reached the computer will merely transform data according to its program. In other words, a classical computer does not have the capability to be selective about its inputs. This is precisely what biological systems can do with ease.

Most proposals on considering brain function to have a quantum basis have done so by default. In short the argument is: There appears to be no resolution to the problem of the binding of patterns and there are non-local aspects to cognition; quantum behaviour has non-local characteristics; so brain behaviour might have a quantum basis.

Newer analysis has led to the understanding that one needs to consider reorganization as a primary process in the brain—this allows the brain to define the context. The signal flows now represent the processing or recognition done within the reorganized hardware. Such a change in perspective can have significant implications. Dual signalling schemes eventually need an explanation in terms of a binding field; they do not solve the basic binding problem in themselves but they do make it easier to understand the process of adaptation.

#### 4. MACHINE AND BIOLOGICAL INTELLIGENCE

For all computational models, the question of the emergence of intelligence is a basic one. Solving a specified problem, that often requires searching or generalization, is taken to be a sign of AI, which is assumed to have an all or none quality. But biological

intelligence has gradation. Animal performance depends crucially on its normal behaviour. It may be argued that all animals are *sufficiently* intelligent because they survive in their ecological environment. Nevertheless, even in cognitive tasks of the kind normally associated with human intelligence, animals may perform well. Thus rats might find their way through a maze, or dolphins may solve logical problems or problems involving some kind of generalization. These performances could, in principle, be used to define a gradation.

If we define thinking in terms of language or picture understanding then, by current evidence, machines cannot think. Machines cannot even perform abstract generalization of the kind that is natural for birds and other animals. But the proponents of strong-AI believe that, notwithstanding their current limitations, machines will eventually be able to simulate the mental behaviour of humans. They suggest that the Turing test<sup>10</sup> should suffice to establish machine intelligence.

We first show that the Turing test is not suitable to determine progress in AI. According to this test the following protocol is used to check if a computer can think: (1) The computer together with a human subject are to communicate, in an impersonal fashion, from a remote location with an interrogator. (2) The human subject answers truthfully while the computer is allowed to lie to try to convince the interrogator that it is the human subject. If in the course of a series of such tests the interrogator is unable to identify the real human subject in any consistent fashion then the computer is deemed to have passed the test of being able to think. It is assumed that the computer is so programmed that it is mimicking the abilities of humans. In other words, it is responding in a manner that does not give away the computer's superior performance at repetitive tasks and numerical calculations.

The asymmetry of the test, where the computer is programmed to lie whereas the human is expected to answer truthfully, is a limitation of the test that has often been criticized. This limitation can be overcome easily if it is postulated that the human can take the assistance of a computer. In other words, one could speak of a contest between a computer and a human assisted by another computer. But this change does not mitigate the ambiguity regarding the kind of problems to be used in the test. The test is not objectively defined; the interrogator is a human.

It has generally been assumed that the tasks that set the human apart from the machine are those that relate to abstract conceptualization best represented by language understanding. The trouble with these popular interpretations of the Turing test, which was true to its intent as best as we can see, is that they focus attention exclusively on the cognitive abilities of humans. So researchers could always claim to be making progress with respect to the ultimate goal of the program, but there was no means to check if the research was on the right track. In other words, the absence of intermediate signposts made it impossible to determine whether the techniques and philosophy used would eventually allow the Turing test to be passed.

In 1950, when Turing's essay appeared in print, matching human reasoning could stand for the goal that machine intelligence should aspire to. The problem with such a goal was that it constituted the ultimate objective and Turing's test did not make an attempt to define gradations of intelligence. Had specific tasks, which would have constituted levels of intelligence or thinking below that of a human, been defined then one would have had a more realistic approach to assessing the progress of AI.

The prestige accorded to the Turing test may be ascribed to the dominant scientific paradigm in 1950 which, following old Cartesian ideas, took only humans to be capable of thought. That Cartesian ideas on thinking and intelligence were wrong has been amply established by the research on nonhuman intelligence of the past few decades.

To appreciate the larger context of scientific discourse at that time, it may be noted that interpretations of quantum mechanics at this time also spoke in terms of observations alone; any talk of any underlying reality was considered outside the domain of science. So an examination of the nature of 'thought', as mediating internal representations that lead to intelligent behaviour, was not considered a suitable scientific subject. Difficulties with the reductionist agenda were not so clear, either in physical sciences or in the study of animal behaviour.

## 5. ANIMAL INTELLIGENCE

For considerable time it was believed that language was essential ground for thought; and this was taken as proof that only humans could think. But nobody will deny that deaf-mutes, who don't have a language, do think.<sup>11</sup> Language is best understood as a subset of a large repertoire of behaviour. Research has now established that animals think and are capable of learning and problem solving.

Since nonhumans do not use abstract language, their thinking is based on discrimination at a variety of levels. If such conceptualization is seen as a result of evolution, it is not necessary that this would have developed in exactly the same manner for all species. Other animals learn concepts nonverbally, so it is hard for humans, as verbal animals, to determine their concepts. It is for this reason that the pigeon has become a favourite for intelligence tests; like humans, it has a highly developed visual system, and we are therefore likely to employ similar cognitive categories. It is to be noted that pigeons and other animals are made to respond in extremely unnatural conditions in Skinner boxes of various kinds. The abilities elicited in research must be taken to be merely suggestive of the intelligence of the animal, and not the limits of it.

In a classic experiment, Herrnstein<sup>12</sup> presented 80 photographic slides of natural scenes to pigeons who were accustomed to pecking at a switch for brief access to food. The scenes were comparable but half contained trees and the rest did not. The tree photographs had full views of single and multiple trees as well as obscure and distant views of a variety of types. The slides were shown in no particular order and the pigeons were rewarded with food if they pecked at the switch in response to a tree slide; otherwise nothing was done. Even before all the slides had been shown the pigeons were able to discriminate between the tree and the non-tree slides. To confirm that this ability, impossible for any machine to match, was not somehow learnt through the long process of evolution and hardwired into the brain of the pigeons, another experiment was designed to check the discriminating ability of pigeons with respect to fish and non-fish scenes and once again the birds had no problem doing so. Over the years it has been shown that pigeons can also distinguish: (i) oak leaves from leaves of other trees, (ii) scenes with or without bodies of water, (iii) pictures showing a particular person from others with no people or different individuals.

Other examples of animal intelligence include myna birds who can recognize trees or people in pictures, and signal their identification by vocal utterances—words—instead of

pecking at buttons, and a parrot who can answer, vocally, questions about shapes and colours of objects, even those not seen before. The intelligence of higher animals, such as apes, elephants, and dolphins is even more remarkable.

Another recent summary of this research is that of Wasserman:<sup>13</sup> '[Experiments] support the conclusion that conceptualization is not unique to human beings. Neither having a human brain nor being able to use language is therefore a precondition for cognition... Complete understanding of neural activity and function must encompass the marvelous abilities of brains other than our own. If it is the business of brains to think and to learn, it should be the business of behavioral neuroscience to provide a full account of that thinking and learning in all animals—human and nonhuman alike.'

An extremely important insight from experiments of animal intelligence is that one can attempt to define different gradations of cognitive function. It is obvious that animals are not as intelligent as humans; likewise, certain animals appear to be more intelligent than others. For example, pigeons did poorly at picking a pattern against two other identical ones, as in picking an A against two B's. This is a very simple task for humans.

Wasserman devised an experiment to show that pigeons could be induced to amalgamate two basic categories into one broader category not defined by any obvious perceptual features. The birds were trained to sort slides into two arbitrary categories, such as the category of cars and people and the category of chairs and flowers. In the second part of this experiment, the pigeons were trained to reassign one of the stimulus classes in each category to a new response key. Next, they were tested to see whether they would generalize the reassignment to the stimulus class withheld during reassignment training. It was found that the average score was 87 per cent in the case of stimuli that had been reassigned and 72 per cent in the case of stimuli that had not been reassigned. This performance, exceeding the level of chance, indicated that perceptually disparate stimuli had amalgamated into a new category. A similar experiment was performed on preschool children. The children's score was 99 per cent for stimuli that had been reassigned and 80 per cent for stimuli that had not been reassigned. In other words, the children's performance was roughly comparable to that of pigeons. Clearly, the performance of adult humans at this task will be superior to that of children or pigeons.

Another interesting experiment related to the abstract concept of sameness. Pigeons were trained to distinguish between arrays composed of a single, repeating icon and arrays composed of 16 different icons chosen out of a library of 32 icons. During training each bird encountered only 16 of the 32 icons; during testing it was presented with arrays made up of the remaining 16 icons. The average score for training stimuli was 83 per cent and the average score for testing stimuli was 71 per cent. These figures show that an abstract concept not related to the actual associations learnt during training had been internalized by the pigeon.

Animal intelligence experiments suggest that one can speak of different styles of solving AI problems. Are the cognitive capabilities of pigeons limited because their style has fundamental limitations? It is possible that the relatively low scores on the sameness test for pigeons can be explained on the basis of wide variability in performance for individual pigeons and the unnatural conditions in which the experiments are performed. But is the cognitive style of all animals similar and do the differences in their cognitive capabilities arise from the differences in the sizes of their mental hardware? Since current machines do not, and cannot, use inner representations, it is right to conclude that their

performance can never match that of animals. Most importantly, the generalization achieved by pigeons and other nonhumans remains beyond the capability of machines.

Donald Griffin expresses the promise of animal intelligence research thus:<sup>14</sup> 'Because mentality is one of the most important capabilities that distinguishes living animals from the rest of the known universe, seeking to understand animal minds is even more exciting and significant than elaborating our picture of inclusive fitness or discovering new molecular mechanisms. Cognitive ethology presents us with one of the supreme scientific challenges of our times, and it calls for our best efforts of critical and imaginative investigation.'

A useful perspective on animal behaviour is its recursive nature, or part-whole hierarchy. Considering this from the bottom up, animal societies have been viewed as *superorganisms*. For example, the ants in an ant colony may be compared to cells, their castes to tissues and organs, the queen and her drones to the generative system, and the exchange of liquid food amongst the colony members to the circulation of blood and lymph. Furthermore, corresponding to morphogenesis in organisms, the ant colony has sociogenesis, which consists of the processes by which the individuals undergo changes in caste and behaviour. Such recursion has been viewed all the way up to the earth itself seen as a living entity. Parenthetically, it may be asked whether the earth itself, as a living but unconscious organism, may not be viewed like the unconscious brain. Paralleling this recursion is the individual who can be viewed as a collection of several *agents* where these agents have sub-agents which are the sensory mechanisms and so on. But these agents are bound together and this binding defines consciousness.

A distinction may be made between simple consciousness and self-consciousness. In the latter, the individual is aware of his awareness. It has been suggested that while all animals may be taken to be conscious, only humans might be self-conscious. It is also supposed that language provides a tool to deal with abstract concepts that opens up the world of mathematical and abstract ideas only to humans.

Edelman<sup>15</sup> suggests that a selection mechanism might be at work that has endowed brains, in their evolutionary ladder, with increasing complexity. But this work does not address the question of holistic computations. From an evolutionary perspective if the fundamental nature of biological computing is different from that of classical computers then models like that of Edelman cannot provide the answers we seek.

## 6. HOLISTIC PROCESSING AND QUANTUM MODELS

The quantum mechanical approach to the study of consciousness has an old history and the creators of quantum theory were amongst the first to suggest it. More recently, scholars have proposed specific quantum-theoretic models of brain function, but there is no single model that has emerged as the favoured one at this point. Arguing for a monistic unity between brain and mind, Pribram summarizes:<sup>16</sup> '[A]nother class of orders lies behind the level of organization we ordinarily perceive...When the potential is actualized, information (the form within) becomes unfolded into its ordinary space-time manifestation; in the other direction, the transformation enfolds and distributes the information much as this is done by the holographic process.'

In my own work I have considered the connections between quantum theory and information for more than thirty years, arguing<sup>17</sup> that brain's processing is organized in

a hierarchy of languages: associative at the bottom, self-organizational in the middle, and quantum at the top. Neural learning is associative and it proceeds to create necessary structures to 'measure' the stimulus-space; at the higher level of multiple agents the response is to reorganize the grosser levels of the neural structure. Each cognitive agent is an abstract quantum system. The linkages amongst the agents are regulated by an appropriate quantum field. This allows the individual at the higher levels of abstraction to initiate cognition or action, leading to active behaviour.

Quantum mechanics is a theory of 'wholes' and in light of the fact that the eye responds to single photons<sup>18</sup>—a quantum-mechanical response—and that the mind perceives itself to be a unity, one would expect that its ideas would be applied to examine the nature of mind and of intelligence. But for several decades the prestige of the reductionist programme of neurophysiology made it unfashionable to follow this path.

Meanwhile, the question of the nature of information, and its observation, has become important in physics. The binding problem of psychology, and the need to postulate a mediating agent in the processing of information in the brain, has also brought the 'self' back into the picture in biology. Oscillations and chaos have been proposed as the mechanisms to explain this binding.

My work has also examined the basic place of information in a quantum-mechanical framework and its connections to structure. This work shows that although many processes that constitute the mind can be understood through the framework of neural networks, there are others that require a holistic basis. Furthermore, if each computation is seen as an operation by a reorganized brain on the signals at hand, this has a parallel with a quantum system where each measurement is represented by an operator. I also suggest that the macrostructure of the brain must be seen as a quantum system.

One striking success of the quantum models is that they provide a resolution to the determinism–free-will problem. According to quantum theory, a system evolves causally until it is observed. The act of observation causes a break in the causal chain. This leads to the notion of a participatory universe.<sup>19</sup> Consciousness provides a break in the strict regime of causality. It would be reasonable to assume that this freedom is associated with all life. But its impact on the ongoing processes will depend on the entropy associated with the break in the causal chain.

## 7. A UNIVERSAL FIELD

Eugene Wigner<sup>20</sup> spoke of one striking analogy between light and consciousness: 'Mechanical objects influence light—otherwise we could not see them—but experiments to demonstrate the effect of light on the motion of mechanical bodies are difficult. It is unlikely that the effect would have been detected had theoretical considerations not suggested its existence, and its manifestation in the phenomenon of light pressure.' He acknowledged one fundamental difference between light and consciousness. Light can interact directly with virtually all material objects whereas consciousness is grounded in a physico-chemical structure. But such a difference disappears if it is supposed that the physico-chemical structure is just the instrumentation that permits observations.

In other words, the notion of a universal field requires acknowledging the emergence of the individual's I-ness at specialized areas of the brain. This I-ness is intimately related

to memories, both short-term and long-term. The recall of these memories may be seen to result from operations by neural networks.

Quantum theory defines knowledge in a relative sense. In the quantum world, it is meaningless to talk of an objective reality. Knowledge is a collection of the observations on the reductions of the wavefunction, brought about by measurements using different kinds of instrumentations.

The indeterminacy of quantum theory does not reside in the microworld alone. For example, the Schrödinger's cat paradox shows how a microscopic uncertainty transforms into a macroscopic uncertainty. Brain processes are not described completely by the neuron firings; one must, additionally, consider their higher order bindings, such as thoughts and abstract concepts, because they, in turn, have an influence on the neuron firings. A wavefunction describing the brain would then include variables for the higher order processes, such as abstract concepts as well. But such a definition will leave a certain indeterminacy in our description.

If we knew the parts completely, one can construct a wavefunction for the whole. But as is well known:<sup>21</sup> 'Maximal knowledge of a total system does not necessarily include total knowledge of all its parts, not even when these are fully separated from each other and at the moment are not influencing each other at all.' In other words, a system may be in a definite state but its parts are not precisely defined.

## 8. STRUCTURE

Consider the distinction between the structures of nonliving and living systems. By the structure of a nonliving system we mean a stable organization of the system. The notion of the stability may be understood from the perspective of energy of the system. Each stable state is an energy minimum.

The structure in a living system is not so easily fixed. We may sketch the following sequence of events: As the environment (the internal and the external) changes, the living system reorganizes itself. This choice, by the nervous system, of one out of a very large number of possibilities, represents the behavioural or cognitive response.

We might view this neural hardware as the classical instrumentation that represents the cognitive act. This might also be viewed as a cognitive agent. Further processing might be carried out by this instrumentation. We may consider the cognitive categories to have a reality in a suitable space.

A living organism must have entropy in its structure equal to the entropy of its environment. If it did not, it would not be able to adapt (respond) to the changing environment.

### *8.1 Principle*

The position of the organism in its ecological environment is determined by the entropy of its information processing system.

This defines a hierarchy. According to this view the universe for an organism shows a complexity and richness corresponding to the complexity of the nervous system. This idea should be contrasted with the anthropic principle where the nature of the universe is explained by the argument that if it was different there would not have been man to

observe it. According to our view, the universe might come to reveal new patterns if we had the capacity to process such information.

It is characteristic of neurophysiology that activity in specific brain structures in given a primary explanatory role. But any determination of the brain structure is impossible. If the brain has  $10^{11}$  neurons and  $10^{14}$  synapses, then even ignoring the gradations in the synaptic behavior, the total number of structures that could, in principle, be chosen exceeds  $(2^{10})^{14}$ , which is greater than current estimates of all elementary particles in the universe.

Assume a system that can exist in only two states. Such a system will find its place where the environment is characterized by just two states. Any structure may be represented by a graph, which may, in turn, be represented by a number, or a binary sequence. Thus in a one dimension, the sequences

00111001, 10001101010, 11000001111

represent three binary-coded structures.

Assume that a neural structure has been represented by a sequence. Since this representation can be done in a variety of ways, the question of a unique representation becomes relevant.

### 8.2 Definition

*Let the shortest binary program that generates the sequence representing the structure be called  $p$ . The idea of the shortest program gives us a measure for the structure that is independent of the coding scheme used for the representation. The length of this program may be taken to be a measure of the information to be associated with the organization of the system. This length will depend on the class of sequences that are being generated by the program. Or in other words, this reflects the properties of the class of structures being considered. In general, the structure  $p$  is a variable with respect to time.*

Assuming, by generalized complementarity, that the structure itself is not defined prior to measurement, then for each state of an energy value  $E$ , we may, in analogy with Heisenberg's uncertainty principle, say that  $\Delta E \Delta t \geq k_1$ , where  $k_1$  is a constant based on the nature of the organizational principle of the neural system.

The external environment changes when the neural system is observed, due to the interference of the observer. This means that as the measurement is made, the structure of the system changes. This also means that at such a fundamental level, a system cannot be associated with a single structure, but rather with a superposition of several structures. This might this be a reason behind pleomorphism, the multiplicity of forms of microbes.

The representation described above may also be employed for the external environment. Let the shortest binary program that generates the external environment be called  $x$ . If the external environment is an eigenstate of the system, then the system organization will not change; otherwise, it will.

We may now propose an uncertainty principle for neural system structure:  $\Delta x \Delta p = k_2$ . This relation says that the environment and the structure cannot be simultaneously fixed. If one of the variables is precisely defined the other becomes uncontrollably large. Either of these two conditions implies the death of the system. Thus, such a system will operate

only within a narrow range of values of the environment and structure. We conjecture that  $k_1 = k_2 = k$ .

## 9. REORGANIZING SIGNALS

Living systems are characterized by continual adaptive organization at various levels. The reorganization is a response to the complex of signal flows within the larger system. For example, the societies of ants or bees may be viewed as single superorganisms. Hormonal and other chemical exchanges among the members of the colony determine the ontogenies of the individuals within the colony. More pronounced than this global exchange is the activity amongst the individuals in cliques or groups.

Paralleling trophallaxis is the exchange of neurotransmitters or electrical impulses within a neural network at one level, and the integration of sensory data, language, and ideas at other levels. An illustration of this is the adaptation of the somatosensory cortex to differential inputs. The cortex enlarges its representation of particular fingers when they are stimulated, and it reduces its representation when the inputs are diminished, such as by limb deafferentation.

Adaptive organization may be a general feature of neural networks and of the neocortex in particular. Biological memory and learning within the cortex may be organized adaptively. While there are many ways of achieving this, nesting among neural networks within the cortex is a key principle in self-organization and adaptation. Nested distributed networks provide a means of orchestrating bottom-up and top-down regulation of complex neural processes operating within and between many levels of structure.

There may be at least two modes of signalling that are important within a nested arrangement of distributed networks. A fast system manifests itself as spatiotemporal patterns of activation among modules of neurons. These patterns flicker and encode correlations that are the signals of the networks within the cortex. They are analogous to the hormonal and chemical exchanges of the ant or bee colonies in the example mentioned earlier. In the brain, the slow mode is mediated by such processes as protein phosphorylation and synaptic plasticity. They are the counterparts of individual ontogenies in the ant or bee colonies. The slow mode is intimately linked to learning and development (i.e., ontogeny), and experience with and adaptation to the environment affect both learning and memory.

By considering the question of adaptive organization in the cortex, our approach is in accordance with the ideas of Gibson<sup>22</sup> who has long argued that biological processing must be seen as an active process. We make the case that nesting among cortical structures provides a framework in which active reorganization can be efficiently and easily carried out. The processes are manifest by at least two different kinds of signalling, with the consequence that the cortex is viewed as a dynamic system at many levels, including the level of brain regions. Consequently, functional anatomy, including the realization of the homunculus in the motor and sensory regions, is also dynamic. The homunculus is an evolving, and not a static representation, in this view.

From a mathematical perspective, nesting topologies contain broken symmetry. A monolithic network represents a symmetric structure, whereas a modular network has preferential structures. The development of new clusters or modules also represents an evolutionary response, and a dual mode signalling may provide a means to define context.

It may also lead to unusual resiliences and vulnerabilities in the face of perturbations. We propose that these properties may have relevance to how nested networks are affected by the physiological process of ageing and the pathological events characterizing some neurobiological disorders.

Reorganization explains the immeasurable variety of the response of brains. This reorganization may be seen as a characteristic which persists at all levels in a biological system. Such reorganization appears to be the basis of biological intelligence.

## 10. ADAPTIVE ORGANIZATION

Active perception can be viewed as adapting to the environment. In the words of Bajcsy:<sup>23</sup> 'It should be axiomatic that perception is not passive, but active. Perceptual activity is exploratory, probing, searching; percepts do not simply fall onto sensors as rain falls onto ground. We do not just see, we look.'

It is not known how appropriate associative modules come into play in response to a stimulus. This is an important open question in neural computing. The paradigm of 'active' processing in the context of memory is usually treated in one of two ways. First, the processing may be pre-set. This is generally termed 'supervised learning', and it is a powerful but limited form of active processing. A second type of processing does not involve an explicit teacher, and this mechanism is termed 'unsupervised learning'. It is sensitive to a number of constraints, including the structure and modulation of the network under consideration.

There are different ways that biological memory may be self-organizing, and in this section, we suggest that the nesting of distributed neural networks within the neocortex is a natural candidate for encoding and transducing memory. Nesting has interesting combinatorial and computational features, and its properties have not been fully examined. The seemingly simplistic organization of nested neural networks may have profound computational properties, in much the same way as recent deoxyribonucleic computers have been put to the task of solving some interesting fundamental problems. However, we do not claim that nesting is the only important feature for adaptive organization in neural systems.

A natural consideration is to examine the structural properties of the forebrain, including the hippocampus and neocortex, which are two key structures in the formation and storage of memory. The hippocampus is phylogenetically an ancient structure, which among other functions, stores explicit memory information. To first approximation, this information is then transferred to the neocortex for long-term storage. Implicit memory cues can access neocortical information directly.

The neocortex is a great expanse of neural tissue that makes up the bulk of the human brain. As in all other species, the human neocortex is made up of neural building blocks. At a rudimentary level, these blocks consist of columns oriented perpendicular to the surface of the cortex. These columns may be seen as organized in the most basic form as minicolumns of about 30  $\mu\text{m}$  in diameter. The minicolumns are, in turn, organized into larger columns of approximately 500–1000  $\mu\text{m}$  in diameter. Mountcastle estimates that the human neocortex contains about 600 million minicolumns and about 600,000 larger columns. Columns are defined by ontogenetic and functional criteria, and there is evidence that columns in different brain regions coalesce into functional

modules.<sup>24</sup> Different regions of the brain have different architectonic properties, and subtle differences in anatomy are associated with differences in function.

The large entities of the brain are ‘composed of replicated *local neural circuits*, modules which vary in cell number, intrinsic connections, and processing mode from one large entity to another but are basically similar within any given entity’. In other words, the neocortex can be seen as several layers of nested networks. Beginning with cortical minicolumns, progressive levels of cortical structure consist of columns, modules, regions and systems. It is assumed that these structures evolve and adapt through the lifespan. It is also assumed that the boundaries between the clusters are plastic: they change slowly due to synaptic modifications or, more rapidly, due to synchronous activity among adjacent clusters.

Results from the study of neural circuits controlling rhythmic behaviour, such as feeding, locomotion, and respiration, show that the same network, through a process of ‘rewiring’ can express different functional capabilities.

In a study of the pattern generator of the pyloric rhythm in lobster, it has been found that the behaviour is controlled by 14 neurons in the stomatogastric ganglion. The predominant means of communication between the neurons is through inhibitory synapses. The reshaping of the output of the network arises from neuromodulation. More than 15 different modulatory neurotransmitters have been identified. These allow the rewiring of the network. Turning on the pyloric suppressors restructures the otherwise three independent networks in the stomatogastric nervous system into a single network, converting the function from regional food processing to coordinated swallowing:<sup>25</sup> ‘Rather than seeing a system as a confederation of neatly packaged neural circuits, each devoted to a specific and separate task, we must now view a system in a much more distributed and fluid context, as an organ that can employ modulatory instructions to assemble subsets of neurons that generate particular behaviors. In other words, single neurons can be called on to satisfy a variety of different functions, which adds an unexpected dimension of flexibility and economy to the design of a central nervous system.’

Consider now the issue of the reorganization of the structure or activity, in response to a stimulus, in more detail. We sketch the following process:

1. The overall architecture of the nested system is determined by the external or internal stimulus, this represents the determination of the connections at the highest hierarchical levels and progresses down in a recursive manner. The learning of the connections in each layer is according to a correlative procedure. The sensory organs adjust to the stable state reached in the nearest level.
2. The deeper layers find the equilibrium state corresponding to the input in terms of attractors. Owing to the ongoing reorganization working in both directions, up to the higher levels as well as toward the lower levels, the attractors may best be labelled as being dynamic.

Superorganisms also have nested structures in terms of individuals who interact more with certain members than others. In the case of ants, the castes provide further ‘modular’ structure.<sup>26</sup> For the case of honeybees:<sup>27</sup> ‘[It is important to recognize] subsystems of communication, or cliques, in which the elements interact more frequently with each other than with other members of the communication system. In context, the dozen or so honeybee workers comprising the queen retinue certainly communicate more within their group (including the queen) than they do with the one or two hundred house

bees receiving nectar loads from foragers returning from the field. The queen retinue forms one communication clique while the forager-receiver bees form another clique.'

The parallel for two distinct pathways of communication is to be seen in superorganisms as well:<sup>28</sup> '[The] superorganism is a self-organizing system incorporating two very distinct pathways of communication. One mode is via localized individual worker interactions with low connectedness, and the other one via volatile or semiochemical pheromones with high connectedness. If we examine the communication modes at the functional level, we see that the pheromones reach the entire superorganism, more or less: a global message with a global reaction.'

Another fundamental communication within the superorganism is the one that defines its constitution. This is a much slower process which can be seen, for example, when a queen ant founds her colony. The queen governs the process of caste morphogenesis.<sup>29</sup> Within the new colony, the queen, having just mated with her suitors and received more than 200 million sperm, shakes off her wings and digs a little nest in the ground, where she now is in a race with time to produce her worker offspring. She raises her first brood of workers by converting her body fat and muscles into energy. She must create a perfectly balanced work force that is the smallest possible in size, yet capable of successful foraging, so that the workers can bring food to her before she starves to death.

The queen produces the workers of the correct size for her initial survival and later, after the colony has started going, she produces a complement of workers of different sizes as well as soldier ants in order to have the right organization for the survival of the colony. When researchers have removed members of a specific caste from an ongoing colony, the queen compensates for this deficit by producing more members of that caste. The communication process behind this remarkable control is not known. The communication mechanisms of the ant or the honeybee superorganisms may be supposed to have analogues in the brain.

## 11. ANOMALOUS ABILITIES AMONGST HUMANS

That cognitive ability cannot be viewed simply as a processing of sensory information by a central intelligence extraction system is confirmed by individuals with anomalous abilities. Idiot savants, or simply savants who have a serious developmental disability or major mental illness, perform spectacularly at certain tasks. Anomalous performance has been noted in the areas of mathematical and calendar calculations; music; art, including painting, drawing or sculpting; mechanical ability; prodigious memory (mnemonism); unusual sensory discrimination or 'extrasensory' perception. The abilities of these savants and of mnemonists cannot be understood in the framework of a monolithic mind.

Oliver Sacks, in his book *The Man Who Mistook His Wife for a Hat*<sup>30</sup> describes two 26-year-old twins, John and Michael, with IQs of 60 who are remarkable at calendrical calculations even though 'they cannot do simple addition or subtraction with any accuracy, and cannot even comprehend what multiplication means'. More impressive is their ability to factor numbers into primes since 'primeness' is an abstract concept. Looking from an evolutionary perspective, it is hard to see how performing abstract numerical calculations related to primes would provide an advantage.

The remarkable observations of the neurosurgeon Wilder Penfield nearly 40 years ago,<sup>31</sup> in which the patients undergoing brain surgery narrated their experience on the

stimulation of the outer layer of the cortex at different points, may be interpreted as showing how the brain works in terms of *gestalts*. The stimulation appeared to evoke vivid memories. Subsequent stimulation of the same site did not necessarily produce the same memory, and stimulation of some other site could evoke the same memory. Furthermore, there was no evidence that these memories represented actual experiences in the patient's past. They had a dreamlike quality, as if they consisted of generic scripts out of which real memories are combined. When the patients heard music they could not generally recall the tune, or they saw individuals whom they could not identify and so on. The events did not appear to have a specific space-time locus.

It appears that generic scripts of this kind taken together form the stuff of real, waking experiences. The workings of the mind may be described in terms of the scripts and their relationships. The architecture of the brain provides clues to the relationships amongst the agents, and this architecture is illuminated by examining deficits in function caused by injury.

## 12. APHASIA, ALEXIA, APRAXIA

One might expect aphasia to be accompanied by a general reduction in the capacity to talk, understand, read, write, as well as do mathematics and remember things. One might also suppose that the ability to read complex technical texts would be affected much more than the capacity to understand simple language and to follow commands.

In reality, the relationship between these capacities is very complex. In aphasia, many of these capacities, by themselves or in groups, can be destroyed or spared in isolation from the others. Historically, several capacities related to language have been examined. These include fluency in conversation, repetition, comprehension of spoken language, word-finding disability, and reading disturbances.

In expressive or Broca's aphasia there is a deficit involving the production of speech. There is deep subcortical pathology as well as damage to the frontal cortex. It is caused by injury to the Broca's area which is located just in front of the primary zone for speech musculature. These speech motor areas are spared in the case of classic Broca's aphasia. When the speech musculature itself is partially paralysed leading to slurred speech that is called *dyarthria*.

In Broca's aphasia speech patterns are reduced to 'content' words and the usage of the simplest, non-inflected forms of verbs. The production of speech is severely impaired but comprehension is relatively intact. Such speech is often telegraphic or agrammatic.

A lesion in the posterior portion of the left temporal lobe, the Wernicke area, causes a receptive aphasia in which the speech production is maintained but comprehension is much more seriously affected. Depending on the extent of damage, it may vary from being slightly odd to completely meaningless.

The Wernicke patient may speak at an abnormally fast pace and augment additional syllables to the end of words or additional words or phrases to the end of sentences. The speech is effortless, the phrase length is normal, and generally there is an acceptable grammatical structure and no problems of either articulation or prosody. But the speech shows a deficiency of meaningful, substantive words, so that despite the torrent of words, ideas are not meaningfully conveyed, a phenomenon called *empty speech*. *Paraphasia* is another characteristic of Wernicke's aphasia. Here words from the same general class

may be inappropriately substituted, or syllables in the wrong order generated, or an utterance produced which is somewhat similar to the correct word. For example, the patient may call a table a 'chair' or an elbow a 'knee' or butter a 'tuber' and so on.

There exist other aphasias such as anomia (with word-finding difficulty), conduction (with good comprehension but difficulty with repetition), and transcortical (with varying degree of comprehension but excellent repetition). In agraphia there is a loss or an impairment of the ability to produce written language.

In alexia, the subject is able to write while unable to read; in alexia combined with agraphia, the subject is unable to write or read while retaining other language faculties; in acalculia, the subject has selective difficulty in dealing with numbers.

Alexia has been known for a long time, but its first clinical description was made over a hundred years ago. One of these patients had suffered a cerebral vascular accident after which he could no longer read. Originally, the patient also suffered from some aphasia and agraphia but the aphasia cleared in due course. The other patient suddenly lost the ability to read but had no other language deficit. This patient, although unable to read except for some individual letters, could write adequately.

Three major varieties of alexia have been described: parietal-temporal, occipital, and frontal. In occipital alexia, there is no accompanying agraphia. In this spectacular condition, there is a serious inability to read contrasted with an almost uncanny preservation of writing ability.

Body movements may be considered an expression of a body language and, therefore, in parallel with aphasia, one would expect to see disorders related to them. Apraxia is the inability to perform certain learned or purposeful movements despite the absence of paralysis or sensory loss. Several types of apraxia have been described in the literature.

In kinetic or motor apraxia there is impairment in the finer movements of one upper extremity, as in holding a pen or placing a letter in an envelope. This is a result of injury in the premotor area of the frontal lobe on the side opposite to the affected side of the body. Kinetic apraxia is thought to be a result of a breakdown in the programme of the motor sequence necessary to execute a certain act.

In ideomotor apraxia the patient is unable to perform certain complex acts on command, although they will be performed spontaneously in appropriate situations. Thus the patient will be unable to mime the act of brushing the teeth although the actual brushing will be easily done. It is believed that this apraxia is caused by the disconnection of the centre of verbal formulation and the motor areas of the frontal lobe.

When the sequence of actions for an act is not performed appropriately, this is called ideational apraxia. The individual movements can be performed correctly but there is difficulty in putting these together. Rather than using a match, the patient may strike the cover of a matchbox with the candle-tip.

Constructional apraxia is the loss in the ability to construct or reproduce figures by assembling or drawing. It seems to be a result of a loss of visual guidance or an impairment in visualizing a manipulative output. This apraxia is a result of a variety of lesions in either one or both of the hemispheres.

The complex manner in which these aphasias manifest establishes that language production is a very intricate process. More specifically, it means that at least certain components of the language functioning process operate in a yes/no fashion.

These components include comprehension, production, repetition, and various abstract processes. But to view each as a separate module only tells half the story. There exist very subtle interrelationships between these capabilities which all come into operation in normal behaviour.

Attempts to find neuroanatomical localization of individual language functions have not been successful. In fact critique of the approach of the localizationists led to a holistic attitude to the brain's function. The anatomical centres, such as the areas of Broca or Wernicke, for the various syndromes are to be viewed as 'focus' areas at a lower level and not exclusive processing centres. The actual centres are defined at some higher levels of abstraction.

### 13. BLINDSIGHT AND AGNOSIA

There are anecdotal accounts of blind people who can see sometimes and deaf people who can likewise hear. Some brain damaged subjects cannot consciously see an object in front of them in certain places within their field of vision, yet when asked to guess if a light had flashed in their region of blindness, the subjects guess right at a probability much above that of chance.

In a typical case the subject is completely blind in the left or right visual field after undergoing brain surgery yet he performs very well in reaching for objects. 'Needless to say, [the patient DB] was questioned repeatedly about his vision in his left-half field, and his most common response was that he saw nothing at all...When he was shown the results, he expressed surprise and insisted several times that he thought he was just guessing. When he was shown a video film of his reaching and judging orientation of lines, he was openly astonished.'<sup>32</sup> Obviously, blindsight patients possess visual ability but it is not part of their conscious awareness.

Blindsight has been explained as being a process similar to that of implicit memory or it has been proposed that consciousness is a result of a dialogue going on between different regions of the brain. When this dialogue is disrupted, even if the sensory signals do reach the brain, the person will not be aware of the stimulus. One may consider that the injury in the brain leading to blindsight causes the vision in the stricken field to become automatic. Then through re-training it might be possible to regain the conscious experience of the images in this field. In the holistic explanation, the conscious awareness is a correlate of the activity in a complex set of regions in the brain. No region can be considered to be producing the function by itself although damage to a specific region will lead to the loss of a corresponding function.

Agnosia is a failure of recognition that is not due to impairment of the sensory input or a general intellectual impairment. A visual agnosic patient will be unable to tell what he is looking at, although it can be demonstrated that the patient can see the object. In visual agnosia the patient is unable to recognize objects for reasons other than that of loss of visual acuity or intellectual impairment. In auditory agnosia the patient with unimpaired hearing fails to recognize or distinguish speech. The patient can read without difficulty, both out loud and for comprehension. If words are presented slowly, the patient may comprehend fairly well; if presented at a normal or rapid speed, the patient will not comprehend. Other patients perceive vowels and/or consonants but not entire words, or some words but not vowels or consonants. These patients have little

difficulty with naming, reading or writing; all language functions except auditory comprehension are performed with ease. Astereognosis is a breakdown in tactile form perception so that the patient cannot recognize familiar objects through touch although the sensations in the hands appear to be normal.

Prosopagnosia literally means a failure to recognize faces. Prosopagnosic patients are neither blind nor intellectually impaired; they can interpret facial expressions and they can recognize their friends and relations by name or voice. Yet they do not recognize specific faces, not even their own in a mirror!

Prosopagnosia may be regarded as the opposite of blindsight. In blindsight there is recognition without awareness, whereas in prosopagnosia there is awareness without recognition. But there is evidence that the two syndromes have underlying similarity. Electrodermal recordings show that the prosopagnosic responds to familiar faces although without awareness of this fact. It appears, therefore, that the patient is subconsciously registering the significance of the faces. Prosopagnosia may be suppressed under conditions of associative priming. Thus if the patient is shown the picture of some other face it may trigger a recognition.

#### 14. SPLIT BRAINS AND UNIFICATION

The two hemispheres of the brain are linked by the rich connections of the corpus callosum. The visual system is arranged so that each eye normally projects to both hemispheres. By cutting the optic-nerve crossing, the chiasm, the remaining fibres in the optic nerve transmit information to the hemisphere on the same side. Visual input to the left eye is sent only to the left hemisphere, and input to the right eye projects only to the right hemisphere. The visual areas also communicate through the corpus callosum. When these fibers are also severed, the patient is left with a split brain.

A classic experiment on cats with split brains was conducted by Ronald Myers and Roger Sperry.<sup>33</sup> They showed that cats with split brains did as well as normal cats when it came to learning the task of discriminating between a circle and a square in order to obtain a food reward, while wearing a patch on one eye. This showed that one half of the brain did as well at the task as both the halves in communication. When the patch was transferred to the other eye, the split-brain cats behaved different from the normal cats, indicating that their previous learning had not been completely transferred to the other half of the brain.

Experiments on split-brain human patients raised questions related to the nature and the seat of consciousness. For example, a patient with left-hemisphere speech does not know what his right hemisphere has seen through the right eye. The information in the right brain is unavailable to the left brain and vice versa. The left brain responds to the stimulus reaching it whereas the right brain responds to its own input. Each half brain learns, remembers, and carries out planned activities. It is as if each half-brain works and functions outside the conscious realm of the other. Such behaviour led Sperry to suggest that there are 'two free wills in one cranial vault'.

But there are other ways of looking at the situation. One may assume that the split-brain patient has lost conscious access to those cognitive functions which are regulated by the non-speech hemisphere. Or, one may say that nothing is changed as far as the awareness of the patient is considered and the cognitions of the right brain were

linguistically isolated all along, even before the commissurotomy was performed. The procedure only disrupts the visual and other cognitive-processing pathways.

The patients themselves seem to support this second view. There seems to be no antagonism in the responses of the two hemispheres and the left hemisphere is able to fit the actions related to the information reaching the right hemisphere in a plausible theory. For example, consider the test where the word 'pink' is flashed to the right hemisphere and the word 'bottle' is flashed to the left. Several bottles of different colours and shapes are placed before the patient and he is asked to choose one. He immediately picks the pink bottle explaining that pink is a nice colour. Although the patient is not consciously aware of the right eye having seen the word 'pink' he, nevertheless, 'feels' that pink is the right choice for the occasion. In this sense, this behaviour is very similar to that of blindsight patients.

The brain has many modular circuits that mediate different functions. Not all of these functions are part of conscious experience. When these modules related to conscious sensations get 'crosswired', this leads to synesthesia. One would expect that similar joining of other cognitions is also possible. A deliberate method of achieving such a transition from many to one is a part of some meditative traditions.

It is significant that patients with disrupted brains never claim to have anything other than a unique awareness. The reductionists opine that consciousness is nothing but the activity in the brain but this is mere semantic play which sheds no light on the problem. If shared activity was all there was to consciousness, then this would have been destroyed or multiplied by commissurotomy. Split brains should then represent two minds just as in freak births with one trunk and two heads we do have two minds.

Consciousness, viewed as a non-material entity characterized by holistic quantum-like theory, becomes more understandable. The various senses are projections of the mindfunction along different directions. Injury to a specific location in the brain destroys the corresponding hardware necessary to reduce the mindfunction in that direction. Mindfunction may be represented along many bases. Instead of aphasias and agnosias, one could have talked of other deficits. The architecture of mind adapts to the environment. This adaptation makes it possible for the mind to compensate.

Gazzaniga has said:<sup>34</sup> 'consciousness is a feeling about specialized capacities.' But why should this feeling of unity persist when the hemispheres are severed? I believe the fact that commissurotomy does not disrupt the cognitive or verbal intelligence of the patients is an argument against reductionism. One must grant that the severed hemispheres maintain a feeling of unity, which manifests as consciousness, by some fundamental field.

The argument that one of the two hemispheres does not have language and that consciousness is uniquely associated with language fails when we consider split-brain patients who had language in both the hemispheres. Gazzaniga suggests that the right hemisphere, although possessing language, is very poor at making simple inferences. He reasons that the two hemispheres have very dissimilar conscious experience. But the fact that both the hemispheres have speech militates against that view. Furthermore, one would expect that the separated hemispheres will start a process of independent reorganization to all the sensory inputs. If the patient still is found to have a single awareness, as has been the case in all tests, then the only conclusion is that the mind remains whole although the brain has been sundered.

## 15. CONCLUSIONS

This article has considered evidence from physical and biological sciences to make the case that machines cannot become conscious. To recapitulate the main points of our paper: We argued that machines fall short on two counts as compared to brains. First, unlike brains, machines do not self-organize in a recursive manner. Second, machines are based on classical logic, whereas nature's intelligence may depend on quantum mechanics.

Quantum mechanics provides us a means of obtaining information about a system associated with various attributes. A quantum state is a linear superposition of its component states. Since the amplitudes are complex numbers, a quantum system cannot be effectively simulated by the Monte Carlo method using random numbers. One cannot run a physical process if its probability amplitude is negative or complex!

The counter-intuitive nature of quantum mechanics arises from the collapse of the state function by the observation. This renders the framework nonlinear, and irreversible if the time-variable is changed in sign. Philosophers of science have agonized over the many bizarre implications of quantum mechanics, e.g. an organism can be both dead and alive before it is observed (Schrödinger's cat paradox), the present can influence the past (Wheeler's delayed-choice scenario), effects can propagate instantaneously in apparent violation of the ceiling of the speed of light (EPR paradox), and so on. The strangeness of quantum mechanics is a consequence of its superpositional logic. But quantum mechanics has characteristics that are intuitively satisfying and it may be interpreted in a manner that allows free will and consciousness.<sup>35</sup>

The evidence from neuroscience that we reviewed showed how specific centres in the brain are dedicated to different cognitive tasks. But these centres do not merely do signal processing: each operates within the universe of its experience so that it is able to generalize individually. This generalization keeps up with new experience and is further related to other cognitive processes in the brain. It is in this manner that cognitive ability is holistic and irreducible to a mechanistic computing algorithm. Viewed differently, each agent is an apparatus that taps into the universal field of consciousness. On the other hand, AI machines based on classical computing principles have a fixed universe of discourse so they are unable to adapt in a flexible manner to a changing universe. This is why they cannot match biological intelligence.

Quantum theory has the potential to provide understanding of certain biological processes not amenable to classical explanation. Take the protein-folding problem. Proteins are sequences of large numbers of amino acids. Once a sequence is established, the protein folds up rapidly into a highly specific three-dimensional structure that determines its function in the organism. It has been estimated that a fast computer applying plausible rules for protein-folding would need  $10^{127}$  years to find the final folded form for even a very short sequence of 100 amino acids.<sup>36</sup> Yet nature solves this problem in a few seconds. Since quantum computing can be exponentially faster than conventional computing, it could very well be the explanation for nature's speed. The anomalous efficiency of other biological optimization processes may provide indirect evidence of underlying quantum processing if no classical explanation is forthcoming.

There are several implications of this work. First, if machines with consciousness are created, they would be living machines, that is, variations on life-forms as we know them.

Second, the material world is not causally closed, and consciousness influences its evolution. Matter and minds complement each other. At the level of the individual, even Western medical science that is strongly based on the machine paradigm has now acknowledged the influence of mind on body.<sup>37</sup> At a more abstract level, it is being argued that even if machines can be conscious, now that this property has emerged through increasing complexity of life-forms, humans will eventually create silicon machines with minds that will slowly spread all over the world, and the entire universe will eventually become a conscious machine.<sup>38</sup> It is most fascinating that starting with the machine paradigm one is led to the ultimate end of a universal mind, whereas in India the traditional view is to postulate a universal mind in the beginning out of which emerges the physical world!

#### NOTES AND REFERENCES

1. S. Kak, 2004, *The Architecture of Knowledge*. Delhi: Centre for Studies in Civilizations/Motilal Banarsidass.
2. W. Schrödinger, 1967, *What is Life? and Mind and Matter*. Cambridge: Cambridge University Press, pp. 128–9.
3. S. Kak, 1996, 'Three languages of the brain: quantum, reorganizational, and associative', in *Learning and Self-Organization*, edited by Karl H. Pribram and Robert King, Lawrence Erlbaum Associates, pp. 185–219; S. Kak, 2000, 'Active agents, intelligence and quantum computing', *Information Sciences*, vol. 128, pp. 1–17.
4. S. Kak, 2002, *The Gods Within*. New Delhi: Munshiram Manoharlal.
5. W.J. Moore, 1992, *Schrödinger: Life and Thought*. Cambridge: Cambridge University Press.
6. S. Kak, 1996, 'Reflections in clouded mirrors: selfhood in animals and machines', in *Learning and Self-Organization*, edited by Karl H. Pribram and Robert King, Lawrence Erlbaum Associates, pp. 511–34.
7. Moore, *Schrödinger*, pp. 170–3.
8. Schrödinger, *What is Life?* For the influence of this book on Crick, see J.D. Watson, *DNA: The Secret of Life*. New York: Knopf.
9. Moore, *Schrödinger*, p. 173.
10. A.M. Turing, 1950, 'Computing machinery and intelligence', *Mind*, 59, 433–60; S. Kak, 1996, 'Can we define levels of artificial intelligence?' *Journal of Intelligent Systems*, vol. 6, pp. 133–44.
11. M. Corballis, 1999, 'The gestural origins of language', *American Scientist*, 87, pp. 138–45; Björn Merker Wallin and Steven Brown (eds.), 2001, *The Origins of Music*. Cambridge: The MIT Press.
12. R.J. Herrnstein, 1985, 'Riddles of natural categorization', *Phil. Trans. R. Soc. Lond.*, B308, 29–144; R.J. Herrnstein, W. Vaughan, Jr., D.B. Mumford, and S.M. Kosslyn, 1989, 'Teaching pigeons an abstract relational rule: insideness', *Perception and Psychophysics*, 46, pp. 56–64.
13. E.A. Wasserman, 1995, 'The conceptual abilities of pigeons', *American Scientist*, 83, pp. 246–55.
14. D. Griffin, 1992, *Animal Minds*. Chicago: The University of Chicago Press.
15. G.M. Edelman, 1992, *Bright Air, Brilliant Fire: On the Matter of the Mind*. Basic Books, New York.
16. K. Pribram, 1987, 'The implicate brain', in *Quantum Implications*, B.J. Hiley and F.D. Peat (eds.). London: Routledge & Kegan Paul.
17. See note 3.
18. D.A. Baylor, T.D. Lamb, and K.-W. Yau, 1979, 'Responses of retinal rods to single photons', *Journal of Physiology*, 288, pp. 613–34.
19. See note 1.
20. E. Wigner, 1961, in *The Scientist Speculates*, I.J. Good (ed.). New York: Basic Books.
21. E. Schrödinger, 1980, 'The present situation in quantum mechanics', *Proc. of the American Philosophical Society*, 124, pp. 323–38.
22. J.J. Gibson, 1979, *The Ecological Approach to Visual Perception*. Boston: Houghton-Mifflin.
23. R. Bajcsy, 1998, 'Active perception', *Proceedings of the IEEE*, 78, pp. 996–1005.
24. V.B. Mountcastle, 1978, 'An organizing principle for cerebral function' in *The Mindful Brain*, G.M. Edelman and V.B. Mountcastle (eds.). Cambridge: The MIT Press.
25. J. Simmers, P. Meyrand, and M. Moulins, 1995, 'Dynamic network of neurons', *American Scientist*, 83, pp. 262–68.

26. J.H. Sudd and N.R. Franks, 1987, *The Behavioural Ecology of Ants*. Glasgow: Blackie.
27. R.F.A. Moritz and E.E. Southwick, 1992, *Bees as Superorganisms*. Berlin: Springer-Verlag p. 145.
28. *Ibid.*, p. 151.
29. M.V. Brian, 1983, *Social Insects*. Chapman and Hall, London. B. Hölldobler, and E.O. Wilson, 1994, *Journey to the Ants*. Cambridge: Harvard University Press.
30. O. Sacks, 1985, *The Man Who Mistook His Wife for a Hat*. HarperCollins, New York.
31. W. Penfield, P. Perot, 1963, 'The brain's record of auditory and visual experience', *Brain*, 86, pp. 595–696.
32. L. Weiskrantz, E.K. Warrington, M.D. Sanders, J. Marshall, 1974, 'Visual capacity in the hemianopic field following a restricted occipital ablation', *Brain*, 97, pp. 709–28.
33. R.E. Myers, R.W. Sperry, 1953, 'Interocular transfer of a visual form discrimination habit in cats after section of the optic chiasm and corpus callosum', *Anatomical Record*, 115, pp. 351–2.
34. M.S. Gazzaniga, (ed.), 1995, *The Cognitive Neurosciences*. Cambridge: The MIT Press, p. 1397.
35. See note 1.
36. A.S. Fraenkel, 1997, 'Protein folding, spin glass and computational complexity', Third Annual DIMACS Workshop on DNA Based Computers, Univ of Pennsylvania. DIMACS Series in Discrete Mathematics and Theoretical Computer Science Vol. 48, AMS, 1999, pp. 101–21.
37. R. Ader, D.L. Felten, and N. Cohen (eds.), 1990, *Psychoneuroimmunology*. New York: Academic Press.
38. F.J. Tipler and J.R. Barrow, 1988, *The Anthropic Cosmological Principle*. London: Oxford University Press.