Consciousness, Whitehead and quantum computation in the brain: Panprotopsychism meets the physics of fundamental spacetime geometry

Stuart Hameroff

I. Introduction: The problem of consciousness and the "emergence" approach

Ever-increasing understanding of brain function has failed to illuminate the nature of consciousness, specificallythe subjective experience of mental states (e.g. Nagel, 1974; Chalmers,1996). Incorporating the conscious mind into a scientific world-view involves finding scientific explanations for what philosophers call "qualia", raw components of subjective experience which give rise to our "inner life". Broadly speaking, there are three types of approaches to the problem of consciousness: *dualism* (consciousness lies outside knowable science), *emergence* (consciousness arises as a novel property from complex computational dynamics in the brain), and some form of *panpsychism*, *pan-protopsychism, or pan-experientialism* (essential features or precursors of consciousness are fundamental components of reality which are accessed by brain processes).

In addition to 1) the problem of subjective experience, other related enigmatic features of consciousness persist, defying technological and philosophical inroads. These include 2) the "binding problem"—how disparate brain activities give rise to a unified sense of "self" or unified conscious content. Temporal synchrony—brain-wide coherence of neural membrane electrical activities—is often assumed to accomplish binding, but *what* is being synchronized? What is being coherently bound? Another enigmatic feature is 3) the transition from pre-conscious processes to consciousness itself. Most neuroscientists agree that consciousness is the "tip of an iceberg", that the vast majority of brain activities is *not* conscious, and that consciousness can occur in brain regions which at other times are not conscious. But there is no explanation for a threshold, or transition from non-conscious, "sub-conscious" or "pre-conscious" processes to consciousness itself. Yet another enigmatic feature is 4) the problem of 'free will. Do we indeed have free will—agency of choice—or are we merely epiphenomenal "helpless spectators", following a deterministic behavioral path shaped by our genes and environment? Finally, 5) what is the nature of subjective time? Does time flow? To our conscious minds time does seem to flow, however in physics there is no necessity for a flow of time. Why is the flow of time such an intrinsic feature of our conscious experience?

Conventional neuroscientific approaches to the enigmatic features of consciousness are based on contemporary understanding of the brain as a collection of neurons acting as fundamental units. A century ago the Spanish neuroanatomist Santiago Ramon-y-Cajal showed the brain to be comprised of individual nerve cells—"neurons"—rather than a woven-together reticulum, as had been suggested earlier by Camille Golgi (Cajal, 1899; Hameroff, 1999). From Cajal's discovery ensued the "neuron doctrine" which implies that the brain operates by interactions among fairly simple individual neuronal units, exchanging signals at connections known as synapses. The development of classical silicon computers which also operate by interactions among fundamental units (e.g. "bits") has fostered the notion that the brain functions in an essentially similar way. Indeed in the 1980's the development of "artificial neural networks", self-organizing computers which can alter connection ("synaptic") strengths and thus "learn", bolstered the analogy between classical computers and the brain. Based on historical comparisons between information processing technology and brain functions (e.g. from the Greeks' "seal ring in wax" as memory, to the telegraph, hologram and silicon computer as the conscious mind) the idea has become prevalent that the brain is a computer.

But what about consciousness? Is consciousness a form of computation? If so, why aren't computers conscious? One answer is dualism, raised by Rene Descartes (consciousness lies outside science) or its modern "mysterianism" counterpart, i.e. Colin McGinn's (e.g. McGinn, 1991) notion that, while consciousness may lie within science, we are not capable of understanding it.

The conventional, most popular answer is "emergence" which suggests that conscious experience "emerges" at a critical threshold of computational complexity among the brain's neurons. Emergence (in its purest form) is a mathematical construct from nonlinear dynamics in which novel properties emerge at thresholds in hierarchical complex systems (Scott, 1995). For example the property of wetness "emerges" from simple interactions among simple water molecules which individually lack "wetness". Weather patterns, including the "Great Red Spot" on Jupiter arise from simple interactions among simple gas molecules and dust particles. Candle flames and self-organizing computer programs are also emergent phenomena. Is consciousness an emergent property of simple synaptic interactions among simple neurons? Perhaps, but "simple" single cell organisms like paramecium can learn, avoid predators, find food and mates and have sex, all without benefit of a synapse (they apparently do through activities of their internal 'cytoskeleton' including microtubules). {footnote 1: Describing the complex behaviors of single cell organisms the famed neuroscientist Charles Sherrington (1953) remarked: "of nerve there is not trace, but the 'cytoskeleton' may serve". Thus was born the idea that the cytoskeleton, normally considered as merely the cell's structural scaffolding also acts as each cell's "nervous system" (e.g. Hameroff and Watt, 1982). end footnote 1}. Furthermore complex weather patterns and computer programs are not conscious as far as we can tell (determining whether a system is conscious is admittedly another problem). Furthermore the emergence view offers no apparent threshold nor testable predictions. In the emergence approach, consciousness has no ontological status setting it apart from non-conscious systems.

II. Whitehead and Pan-protopsychism

An alternative set of philosophical positions which does ascribe ontological status to consciousness, or its precursors, includes panpsychism, pan-protopsychism, and pan-experientialism. These view consciousness as being related to a "funda-mental", irreducible component of physical reality, something like mass, spin or charge. These components just *are*.

Panpsychism stems from ancient Greek philosophers—e.g. Democritus' "atomism" — and holds that primitive, dim consciousness is a quality of all matter: atoms and their subatomic components having subjective, mental attributes (e.g. Spinoza, 1677; Rensch, 1960). "Mentalists" such as Leibniz (1768) and Whitehead (1929;1933) contended that systems ordinarily considered to be physical are constructed in some sense from more basic mental entities. Bertrand Russell (1954) described "neutral monism" in which a common underlying entity, neither physical nor mental, gave rise to both. Bishop Berkeley's "idealism" suggested that consciousness creates reality, that consciousness is “all there is”. Wheeler (1990) has suggested that information is fundamental to the physics of the universe, and from this Chalmers (1996) proposes a double-aspect theory in which information has both physical and experiential aspects. Ascribing features of conscious experience to fundamental reality raises two new questions: 1) what IS fundamental reality (or fundamental information) e.g. as describable by modern physics, and 2) how are conscious and non-conscious systems different? Among philosophical approaches, the pan-protopsychist or pan-experiential philosophy of Alfred North Whitehead (1929; 1933) seems best suited to connect consciousness to the physics of reality.

Whitehead viewed the universe as being comprised not of things, but of events—as a process. Leibniz (1768) had “quantized” reality, describing fundamental “monads” as the ultimate entities of reality but Whitehead transformed monads into “actual occasions” occurring in a “basic field of proto-conscious experience”. Whitehead occasions are spatio-temporal quanta, each endowed—usually on a very low level—with mentalistic characteristics like “experience, subjective immediacy, appetition”. However the experience of each fundamental occasion is “dull, monotonous, and repetitious” (and thus not noticeable to physical observation). Whitehead viewed our high level mentality, consciousness, as being extrapolated (?”emerging”) from temporal chains of occasions. In his view highly organized societies of occasions permit primitive mentality to become intense, coherent and fully conscious. Meanwhile “the functionings of inorganic matter remain intact amid the functionings of living matter. It seems that, in bodies that are obviously living, a coordination has been achieved that raises into prominence some functions inherent in the ultimate occasions.”

Abner Shimony (1993; 1997) recognized that Whitehead’s approach was potentially compatible with modern physics, specifically quantum theory, with quantum state reductions—actual events—appearing to represent “occasions”. Quantum theory is a description of reality at small scales. To appreciate Shimony’s “modern Whiteheadianism” and bridge the gap between philosophy and physics for a scientific description of consciousness we must first attempt to come to grips with the unsettling features of quantum theory.

III. Reality: The quantum and classical worlds

At first glance, at least, reality (like consciousness) appears dualistic. In our everyday “classical” world, matter and energy are predictable and well behaved, following Newton’s laws of motion and Maxwell’s equations for electromagnetics. However at small e.g. atomic scales governed by quantum theory everything changes and behaviors are so strange that the American physicist Richard Feynman once commented "anyone who claims to understand quantum theory is either lying or crazy."

In the quantum realm (and the boundary between the quantum and classical worlds remains mysterious) objects may exist in two or more states or places simultaneously—more like waves than particles and governed by a “quantum wave function”. This property of multiple coexisting possibilities, known as quantum superposition, persists until the superposition is measured, observed or interacts with the classical world or environment. Only then does the superposition of multiple possibilities “reduce”, “collapse”, “actualize”, “choose” or “decohere” to specific, particular classical states.

Early experiments seemed to show that even if a machine measured a quantum superposition, the multiple possibilities persisted until the machine's results were observed by a conscious human. This led leading quantum theorists including Bohr, Heisenberg and Wigner to conclude that *consciousness* caused quantum state reduction, that consciousness "collapsed the wave function" (the "Copenhagen interpretation", reflecting the Danish origin of Nils Bohr, its leading proponent).

To illustrate the apparent absurdity of this conclusion, in the 1930's Schrödinger devised his famous thought experiment known as Schrödinger's cat. A living cat is placed in a closed box into which poison can be released by a quantum event, e.g., sending a photon through a half-silvered mirror. Being a quantum entity, the (unobserved) photon must be in superposition of both passing through and not passing through the mirror, thus both releasing and not releasing the poison. Consequently, according to the Copenhagen interpretation, until a conscious being opens the box to observe, the cat is both dead *and* alive. Schrödinger's point was that this scenario was absurd and that the conscious observer interpretation was incorrect. Modern interpretations would say that any interaction with the environment (i.e. opening the box regardless of whether a conscious observer was present) would "decohere" the quantum superposition. Nonetheless the fate of an isolated quantum superposition, for example an isolated large scale system evolving from, or amplified by, a small scale superposition, remains unknown. Thus events in the quantum realm are not only bizarre, the boundary between the quantum realm and our everyday macroscopic "classical" world remains obscure.

Another quantum property is entanglement, or quantum coherence, in which components of a system become unified, governed by one common quantum wave function. If one member of an entangled system is measured or perturbed, other members are instantaneously affected, even over great distances. One example of entanglement is the famous "EPR pairs" (after Einstein, Podolsky and Rosen who posed the problem as a thought experiment in the 1930's). Imagine two members of a quantum system (e.g. two electrons with complementary spin: if one is spin up, the other is spin down, and vice versa). If the paired electrons (both in superposition of both spin up and spin down) are separated by being sent along different wires, say to two different villages miles apart from each other, they each remain in superposition. However when one superpositioned electron is measured by a detector at its destination and reduces/collapses to a particular spin, (say spin up), its entangled twin miles away *instantaneously* reduces/collapses to the complementary spin down. The experiment has been done repeatedly with electron spin pairs, polarized photons and other quantum systems and always results in instantaneous reduction to the complementary classical state (e.g. Aspect et al, 1982). The instantaneous, faster than light coupling, or "entanglement" remains unexplained.

Another form of entanglement occurs in quantum coherent systems such as Bose-Einstein condensates (proposed by Bose and Einstein decades ago but realized in the 1990's). A group of atoms or molecules are brought into a quantum coherent state such that they surrender individual identity and behave like one quantum system, marching in step and governed by one quantum wave function. If one component is perturbed all components "feel" it and react accordingly.

Despite the mystery, quantum superpositions and entanglements are used technologically in quantum computers which promise to revolutionize information processing (and perhaps make comparisons between the brain/mind and quantum computers inevitable). Conventional classical computers represent digital information as "bits" of either 1 or 0. In quantum computers information may be represented as quantum superpositions of both 1 *and* 0 (quantum bits, or "qubits"). While in superposition qubits interact with other entangled qubits allowing computational interactions of enormous speed and near-infinite parallelism. After the computation is performed the qubits are reduced (e.g. by environmental interaction/decoherence) to specific classical states which constitute the solution. Although the decoherence introduces some probabilistic randomness, this may be overcome by parallel redundancy.

Quantum computing (and related quantum information technologies quantum cryptography and quantum teleportation) promise to revolutionize information processing and other aspects of society. However the underlying mechanisms remain unknown. What does it actually mean for an object to be in two or more places or states simultaneously? How can nonlocal entanglement occur? What happens to isolated quantum superpositions? Seeking a deeper understanding, as well as a mechanism for consciousness, British mathematical physicist Sir Roger Penrose has addressed these issues.

IV. Penrose "objective reductions": Whitehead "occasions" in fundamental spacetime geometry?

The puzzle of quantum superposition has baffled physics. Although modern physics emphasizes "decoherence" in which any interaction with environment causes loss of quantum superposition, the fate of isolated superpositions (e.g. Schrödinger's cat) remains unresolved. One solution was put forth by Hugh Everett in his "multiple worlds" view (e.g. Everett, 1973). Everett's idea was that superposition is a separation in underlying reality, that the universe at it's fundamental level splits, or separates, and that each possibility branches off to form a new universe, a new reality. Thus, according to this view, there exist an infinite number of parallel universes. For the Schrödinger's cat story, each time the box is opened the universe bifurcates into one universe with a live cat, and another universe with a dead cat. Assuming for a moment that the multiple worlds view is at least partially correct (and many believe it *is* correct), how do we envision the universe separating? How do we envision the *fabric* of reality?

According to modern physics, reality is rooted in 3-dimensional space and a 1-dimensional time, combined together into a 4-dimensional spacetime. For simplicity and illustration we can condense our 4-dimensional spacetime into a 2-dimensional spacetime sheet: one spatial dimension and one time dimension (Figure 1, top). This spacetime is slightly curved, in accordance with Einstein's general theory of relativity, in a way which encodes the gravitational fields of all distributions of mass density. Each mass density—each object or particle—effects a spacetime curvature, albeit tiny for small objects. {Footnote 2: Many are familiar with the idea of large objects causing large spacetime curvature. Einstein had predicted that the spacetime curvature of our sun would bend light from stars, distorting their perceived position e.g. in special cases making them visible when in fact they lay behind the sun from our vantage point. Some 50 years after this prediction, Sir Arthur Eddington made the critical observations during a solar eclipse to prove Einstein's hypothesis. However the idea of small, quantum objects causing small spacetime curvatures was put forth by Penrose. End footnote 2} Consequently we can view any mass in one location as spacetime curvature in a particular direction, and location of the mass in a different location as spacetime curvature in another direction. Therefore quantum superposition of a particle in two locations may be considered simultaneous curvatures in opposite directions (Penrose, 1989; 1994). As in the "multiple worlds" view, the spacetime sheet separates into two opposing curvatures, resulting in a "bubble" or "blister" in underlying reality (Figure 1, bottom).

Fig01.tif

Figure 1. *4-dimensional spacetime may be simplified as a 2-dimensional spacetime sheet, with time running vertically along the y-axis.Top: two mass locations correspond with two spacetime curvature histories, e.g. one particular location is represented as curvature into the plane of the spacetime sheet (left) and another particular location is represented as curvature out of the plane of the spacetime sheet (right). Arrow indicates classical movement of the mass between the two locations. Bottom: quantum superposition of both mass locations represented as simultaneous curvature in both directions—a separation, bubble or blister of spacetime geometry (adapted from Penrose, 1994 - p. 338).*

{Footnote 3: Strictly speaking the separations cannot be considered to have any true "width", or "length" as spacetime defines its dimensions, rather than exists in dimensions. However metaphorically we can consider that the distance between the separated spacetimes ("width") is on the order of a Planck length (10^-33 centimeters) whereas the length may be macroscopic, on the order of the mass separation distance, e.g. nanometers (10^-8 centimeters) or larger, or the distance over which mass separation distance occurs (e.g. ~10 centimeters, as may occur in the brain-see next section). That such "narrow" separations have significant consequences may seem surprising, however an analogy may be drawn to earthquakes in which the earth separates only slightly, but over a great "length" or faultline with significant consequences. end footnote 3}

What is the fate of superpositions/spacetime separations? In the multiple worlds view each spacetime sheet—each side of the "blister"—evolves into a separate universe. However in Penrose's view these separations, bubbles, or blisters are unstable, and somewhat like bubbles in a bubble bath will eventually reduce, or collapse to one particular curvature or the other. The instability is inherent in the properties of spacetime geometry (quantum gravity) and constitutes an *objective* threshold for an isolated quantum state reduction, hence "objective reduction (OR)".

In the Penrose formulation, objective reduction due to the quantum gravity properties of fundamental spacetime geometry occurs at a time T given by the Heisenberg indeterminacy principle E=?/T, in which E is the magnitude of superposition/separation, ? is Planck's constant over 2π, and T is the time until reduction. The magnitude E is related to the gravitational self-energy of the superposition and may be calculated from the amount of mass "separated from itself" and distance of separation. Since E is inversely related to T, small separations/superpositions (if isolated) will reduce at a long time T, and large separations/superpositions (if isolated) will reduce quickly. For example an isolated superpositioned electron would reduce only after 10 million years. A large object, such as a one kilogram cat, would reduce after only 10^-37 seconds (too quickly for anyone to notice).

The point is that Penrose objective reductions are actual events occurring at the level of—or actually in the medium of—fundamental spacetime geometry, and accordingly could qualify for Whitehead occasions, as suggested by Abner Shimony. This is the key point I wish to make: *Penrose OR events may be equivalent to Whitehead "occasions of experience"* (Hameroff and Penrose, 1996b; Hameroff, 1998). But why would these events, or occasions, be conscious?

Whitehead proposed that consciousness was a series of "occasions" occurring in a "wider field of proto-conscious experience". This raises the question of whether the fundamental level of spacetime geometry, the basement level of the universe, *is* that wider field. Could proto-conscious entities—"qualia"—exist in fundamental spacetime geometry? The 2-dimensional spacetime sheets are merely cartoons of reality. What *is* fundamental spacetime geometry?

Democritus described empty space as a true void whereas Aristotle saw a background "plenum" filled with substance. Maxwell's 19th-century "luminiferous ether" sided with Aristotle but attempts to detect the ether failed and Einstein's special relativity suggested that there was no background pattern or structure. However Einstein's general relativity related mass to curvature in a geometric spacetime "metric" and swung the pendulum back to Aristotle's view of an underlying pattern. But where is it? What is it? Most of our universe is "empty", devoid of mass. Atomic nuclei and electrons occupy only a small fraction of an atom's volume—most of an atom is empty space. What is empty space?

Scale.tif

Figure 2. Empty space from a protein to the Planck scale: a) a tubulin protein is on the order of 10^-8 centimeters (nanometers), b) an atom is on the order of 10^-9 centimeters (angstroms), c) below the level of atoms and subatomic particles is the "vacuum", i.e. empty space at 10-25 centimeters, d) the vacuum at 10-30 centimeters: some granularity is apparent, e) nearing the Planck scale at 10-32 centimeters: an "aerial view" of spin networks, f) the Planck scale at 10^-33 centimeters: spin networks describe Planck scale volumes, or "pixels" of reality.

At very small scales spacetime is not smooth, but quantized. Imagine viewing the ocean from an airplane. The ocean surface may look perfectly smooth. However if you were in a small boat on the ocean surface you'd be tossed about by the roughness of the sea invisible from high above. Similarly as we go down in scale from the size of atoms (10^-8 centimeters) empty space seems smooth until eventually we find granularity at the incredibly small "Planck scale" (10^-33 centimeters , 10^-43 seconds; Figure 2). There are several types of descriptions of the Planck scale, for example string theory, the "quantum foam", and loop quantum gravity. In the context of loop quantum gravity, Penrose (1971) portrayed the Planck scale as a dynamical spider-web of spin. Taking spin as an irreducible, fundamental entity, spin networks define spectra of discrete Planck scale volumes and configurations which dynamically evolve and define spacetime geometry. (Figure 3: Rovelli and Smolin, 1995a; 1995b; Smolin, 1997). The amount of potential information in Planck scale spin networks is vast; each Planck scale volume, or "pixel of reality" may be shaped by a huge variety of combinations of "edge" lengths, number of spins per edge, and nonlocal interactions. In addition to the enormous potential variety in each Planck scale pixel, their sheer number compared to our macroscopic scale is enormous—there are roughly 10^107 Planck volumes or pixels in the volume of a human brain, far greater than the number of particles in the universe.

Fig02.tif

Figure 3.A spin network. Introduced by Roger Penrose (1971) as a quantum mechanical description of the geometry of space, spin networks describe a spectrum of discrete Planck scale volumes and configurations (with permission from Smolin, 1997; Rovelli and Smolin, 1995a; 1995b). Average length of each edge is the Planck length (10^-33 cm). Numbers indicate quantum mechanical spin along each edge. Each quantum state of spacetime is a particular spin network (Smolin, 1997).

So the universe may be constructed of Planck scale spin networks whose configurations and dynamics lead to all matter and energy. If consciousness derives from fundamental, irreducible entities (e.g. "proto-conscious qualia") then proto-conscious qualia must also be embedded in Planck scale spin networks (where else *could* they be embedded; fundamental spacetime geometry is *all there is!*). As events occurring in this medium, Penrose "OR" objective reductions are thus compatible with Whitehead "occasions of experience" occurring in a "wider field of protoconscious experience" (Hameroff and Penrose, 1996b).

When quantum superpositions reduce by interactions with the environment ("decoherence") the choice of the particular classical state is probabilistic (a feature with which Einstein famously disagreed: "God does not play dice with the universe"). However following Penrose OR, the particular choice of Planck scale geometry is said to be influenced by "Platonic" information embedded in the fine grain of spacetime geometry, information presumably embedded there since the Big Bang. Because of this influence these types of reductions are neither algorithmic nor random, but what Penrose terms "non-computable", a property he also ascribes to conscious thought. Thus each OR event selects a particular set of proto-conscious qualia, or volitional act, influenced by Platonic values.

V. Could Penrose "objective reductions" (Whitehead "occasions of experience") occur in the brain?

If we equate Penrose objective reduction (OR) in fundamental spacetime geometry with Whitehead occasions of experience (and Buddhist “moments of experience”), then a sequence of OR events/conscious moments occurring in the brain could give rise to our familiar “stream” of consciousness.

Quantum computers utilize quantum superpositions/spacetime separations, however reduction in such devices occurs by decoherence (loss of isolation) before the OR threshold may be met. {Footnote 4. In quantum computational devices the superpositions are generally of electrons, atoms or photons, hence of very low mass and thus very low gravitational self-energy E. According to E=?/T, for small values of E, T would be very long, perhaps years. Since the technological quantum computations are fast, the superposition is interrupted so that the solution may be obtained. Consequently such quantum computers will not be conscious by the OR criteria, at least in the foreseeable future. End footnote 4}. Thus consciousness requires fairly stringent conditions: superpositions/spacetime separations must be large enough to reach threshold in a brief time period (i.e. on the order of brain processes less than one second), yet able to be isolated/protected from disruption by environmental decoherence.

Technological quantum computers require extreme cold (near absolute zero) and strict isolation to avoid decoherence by thermal vibrations during brief periods of superposition, yet the brain operates at a very warm 37.6 degrees centigrade. Consequently quantum computation in the brain has appeared highly unlikely to most scientists. On the other hand biology has had billion of years to solve these problems and quantum computation and consciousness would certainly be beneficial from an evolutionary standpoint.

Where/how could quantum computation involving isolated quantum superpositions of information (“qubits”) occur in the brain and be linked to known brain processes? The Penrose-Hameroff model of “orchestrated objective reduction (Orch OR)” proposes that quantum computations occur in cytoskeletal microtubules within the brain’s neurons, isolated and shielded from environmental decoherence by a variety of evolutionary adaptations (Penrose and Hameroff, 1995; Hameroff and Penrose 1996a; 1996b; Hagan et al, 2002).
The Orch OR paper.gif
Figure 4. Interiors of neurons and other cells are structured by networks of microtubules (MTs) interconnected by microtubule-associated proteins (MAPs), as shown in this immunofluorescent micrograph. MTs are arrayed horizontally, connected by vertical MAPs. Scale bar (lower right): 100 nanometers). With permission from Hirokawa (1991).
Fig05.tif
Figure 5.Left: Microtubule (MT) structure: a hollow tube of 25 nanometers diameter, consisting of 13 columns of tubulin dimers arranged in a skewed hexagonal lattice (Penrose, 1994). Right (top): Each tubulin molecule may switch between two (or more) conformations, coupled to London forces in a hydrophobic pocket. Right (bottom): Each tubulin can also exist in quantum superposition of both conformational states (Hameroff and Penrose, 1996a).

Microtubules are main components of each cell's cytoskeleton (Figure 4), originally thought to provide only "bone-like" structural support but now appreciated also as the cell's information processing system, or "on-board computer". Microtubules (MTs) are polymers of the protein tubulin, arranged in a skewed hexagonal lattice wrapped in a hollow cylinder (Figure 5). Each peanut shaped tubulin may switch between two or more structural conformations, governed by quantum mechanical forces within the protein interior and interact with neighboring tubulins to account for information processing and signaling. The switching occurs rapidly, in the nanosecond (10^-9 sec) scale, and seems to be coherently driven by metabolic energy so that MTs behave somewhat like lasers. There are roughly 10^7 tubulins/neuron switching at ~10^9/sec, so the potential information processing capability of a single neuron at the microtubule level is roughly 10^16 operations/second, as much as is suggested for the entire brain operating at the level of neuronal synaptic switching (Rasmussen et al, 1990). This capacity can account for the complex behavior of single cell organisms, and enhances the projected capacity of the brain enormously (and "raises the bar" for conventional artificial intelligence "AI" attempts to simulate the brain, e.g. Moravec, 1987) but fails to account for the enigmatic features of consciousness.

An essential feature of the Penrose-Hameroff Orch OR model is that tubulins become quantum superpositions of both conformations, and function as qubits by interacting nonlocally (entangling) with other tubulin qubits so that MTs act as quantum computers (Figure 6). When enough entangled tubulins are superpositioned long enough to reach OR threshold by E=?/T, a conscious event/Whitehead occasion of experience occurs. The classical tubulin states chosen in the OR event proceed to regulate classical neural activities, e.g. trigger axonal membrane action potentials, adjust synaptic strengths and rearrange the cytoskeleton, thus exerting causal efficacy, learning and memory.
Fig07.tif
Figure 6.An Orch OR event/Whitehead occasion. a) Microtubule simulation in which classical computing (step 1) leads to emergence of quantum coherent superposition (and quantum computing (steps 2 & 3) in certain (gray) tubulins. Step 3 (in coherence with other microtubule tublins) meets critical threshold related to quantum gravity for self-collapse (Orch OR). A conscious event (Orch OR) occurs in the step 3 to 4 transition. Tubulin states in step 4 are noncomputably chosen in the OR collapse, and evolve by classical computing to regulate neural function. b) Schematic graph of proposed quantum coherence (number of tubulins) emerging versus time in microtubules. Area under curve connects superposed mass energy E with collapse time T in accordance with E=?/T. E may be expressed as Nt, the number of tubulins whose mass separation (and separation of underlying space time) for time T will self collapse. For T = 25 msec (e.g. 40 Hz oscillations), Nt = 2 x 1010 tubulins.
MTs are interconnected into networks by linking proteins called microtubule-associated proteins ("MAPs"). These MAPs attach at specific sites on the MT lattice, and the pattern of attachment determines MT network properties as well as cell shape, movement and function. The MAP attachments and network properties can dynamically change, for example to alter neuronal synaptic strength in learning. MT-MAP networks are very much like "neural networks", but at smaller scale, occurring within each neuron

of a neural network (something like a “fractal” sub-dimension of neural networks, a “forest within each tree”).

Missing – was pore.tif

Figure 7. Quantum coherent superposition of microtubules “orchestrated” by microtubule-associated proteins (MAPs) in two neuronal dendrites connected by a gap junction. Dendritic lamellar bodies attached to MTs may facilitate quantum tunneling, entanglement and spread of unified quantum state between neurons, and throughout large regions of brain.

Simulations of MT dynamics show that MAPs tend to attach at resonant “nodes” on the MT surface (Samsonovich et al, 1992), thus MAPs may tune MT quantum activities (somewhat like frets in a guitar), providing feedback and “orchestrating” MT quantum computations (thus the model of “Orchestrated objective reduction - Orch OR”; Figure 7).

MTs whose tubulins are in quantum superposition in a particular neuron may entangle with those in other neurons via quantum tunneling across window-like “gap junctions” between neurons. Thus brain-wide entangled quantum states can occur within neurons whose interiors are continuous through gap junctions (Woolf and Hameroff, 2001; Figure 7).

Brain processes occur in time scales on the order of tens to hundreds of milliseconds. For example sensory responses are on the order of up to 500 milliseconds (1/2 second), alpha EEG is roughly 100 milliseconds (1/10 second), and “coherent 40 Hz”, the brain-wide synchrony which seems to correlate with conscious activity, is on the order of 25 milliseconds (1/40 second). For OR/Whitehead events in the brain to correspond with known neural events we can use E=?/T and assume T is on the order of such events. For T=25 milliseconds (coherent 40 Hz), we can calculate E in terms of number of tubulins, and estimating for percentage of tubulins/neuron involved in consciousness, find that 10,000 to 100,000 neurons are involved in each OR/Whitehead/conscious event which occur 40 times/second.

Fig08.tif

Figure 8.An Orch OR event (continued from Figure 6. a) (left) Three tubulins in quantum superposition prior to 25 msec Orch OR. After reduction (right), particular classical states are selected. b) Fundamental spacetime geometry view. Prior to Orch OR (left), spacetime corresponding with three superposed tubulins is separated as Planck scale bubbles: curvatures in opposite directions. The Planck scale spacetime separations S are very tiny in ordinary terms, but relatively large mass movements (e.g., hundreds of tubulin conformations, each moving from 10^6 to 0.2 nm) indeed have precisely such very tiny effects on the spacetime curvature. A critical degree of separation causes Orch OR and an abrupt selection of single curvatures (and a particular geometry of experience). c) Cognitive facial recognition. A familiar face induces superposition (left) of three possible solutions (Amy, Betty, Carol) which "collapse" to the correct answer Carol (right). d) Cognitive volition. Three possible dinner selections (shrimp, sushi, pasta) are considered in superposition (left), and collapse via Orch OR to choice of sushi (right).

Each OR event is instantaneous, so the 25 milliseconds/conscious events are in the pre-conscious quantum superposition phase of multiple possibilities of perceptions or choices. For example imagine you briefly see a woman's face, is it Amy, Betty, or Carol (Figure 8)? During the pre-conscious superposition phase there are quantum superpositions of all 3 possibilities which then reduce/collapse/choose one particular possibility at the moment of OR ("Aha, it is Carol!"). Or you are studying a menu deciding what to order for dinner, and have e.g. superpositions of shrimp, sushi or pasta. After time T you reduce/collapse/choose and decide "I'll have sushi"! The pre-conscious superposition phase may also be equated with the Freudian sub-conscious including dreams and perhaps altered states.
Quantum_coherence.tif
Figure 9. Quantum superposition/entanglement in microtubules for 5 states related to consciousness. Area under each curve equivalent in all cases. A. Normal 40 Hz experience: as in Figures 6 & 8.. B. Anesthesia: anesthetics bind in hydrophobic pockets of brain proteins and prevent quantum delocalizability and coherent superposition. C. Heightened Experience: increased sensory experience input (for example) increases rate of emergence of quantum superposition. Orch OR threshold is reached faster, and Orch OR frequency increases. D. Altered State: even greater rate of emergence of quantum superposition due to sensory input and other factors promoting quantum state (e.g. meditation, psychedelic drug etc.). Predisposition to quantum state results in baseline shift and collapse so that conscious experience merges with normally sub-conscious quantum computing mode. E. Dreaming: prolonged sub-threshold quantum superposition time.

Both Whitehead and the quantum approach suggest that consciousness is a "stream" of discrete events, rather than a continuous state (Figure 9). Obviously we perceive our world as continuous rather than as discrete events, but a movie appears continuous though it is in fact a sequence of frames. Of course a movie has an external conscious observer, whereas in the Whitehead/Orch OR quantum approach each OR (self-collapse) event *is* consciousness.

VI. How can the Whitehead/quantum approach account for the enigmatic features of consciousness?

1) The "hard problem" of subjective experience, or qualia is accounted for by ascribing proto-conscious qualia to properties of fundamental spacetime geometry, e.g. as particular configurations of Planck scale spin networks. Each OR event selects a particular configuration from among multiple possibilities of spacetime-embedded qualia, somewhat like an artist may select fundamental colors from a palette to create a complex painting. A rose is a particular pattern of spacetime geometry, both the spacetime which *is* the rose, and the same pattern recreated in the brain (Figure 10).
Rose_spacetime.tif
Figure 10. Are mental 'qualia' like the redness of a rose fundamental patterns in spacetime geometry?

2) Binding, or unity of conscious experience is accomplished by quantum coherence/entanglement (Marshall, 1989). Proto-conscious qualia selected in a single OR event are governed by one wave function; they are essentially one common entity.

3) The transition from pre-conscious processes to consciousness itself is the collapse, or reduction inherent in OR—a Whitehead spatiotemporal quantum with the threshold given by E=?/T.

4) The problem of free will relates to the issue of whether our actions are completely deterministic, algorithmic and/or random. Penrose non-computability provides an alternative; our actions are algorithmic/deterministic but with an added ingredient—influence by Platonic values embedded in fundamental spacetime geometry. Our experience of free will is that of our deterministic processes influenced by these Platonic influences, not truly "free", but augmented by unseen forces (Figure 11).
Freewill.tif
Figure 11. Free will may be seen as the result of deterministic processes (behavior of trained robot windsurfer) acted on repeatedly by non-computable influences, here represented as a seemingly capricious wind.

5) Subjective time flow is a function of discrete, irreversible events "ratcheting" forward in time. As we become excited we experience more conscious events per "clock time", thus the outside world seems to "slow down". If our consciousness is partially suppressed, say we've had too much alcohol to drink, we have fewer conscious events per clock time and the outside world appears faster ("don't drink and drive!"). In the case of general anesthesia (which prevents quantum interactions in key brain proteins) subjective time ceases entirely.
Several lines of evidence suggest that the brain can project information "backwards in time", i.e. from the near future to the present (Libet et al, 1979; Bierman and Scholte, 2002). In the quantum world time is also indeterminate, and quantum state reductions may send quantum information both forwards and backwards (Aharonov and Vaidman, 1990). This may account not only for "precognition and premonitions", but also play a role in day-to-day, moment-to-moment activities in which we may act, then decide only slightly afterwards how to do so.

VII. Conclusion

Explanations for consciousness have traditionally fallen along separate lines of either philosophy, neuroscience, psychology or physics. Whitehead's philosophical approach is consistent with modern physics, and only such a union of philosophy and science can account for consciousness. Specifically Whitehead "occasions" in a "wider field of proto-conscious experience" may translate to Penrose "objective reductions" in pan-protopsychist fundamental spacetime geometry. The Penrose-Hameroff "Orch OR" model incorporates neuroscience and psychology along with physics and philosophy to produce a complete theory. The theory may be wrong, but it is testable and hence falsifiable—the type of comprehensive unified theory which will be required to explain the enigmatic features of consciousness.

Acknowledgements: Supported by the Fetzer Institute, YeTaDeL Foundation, and The Center for Consciousness Studies and Departments of Anesthesiology and Psychology at the University of Arizona. I am grateful to Roger Penrose for his insight and collaboration and to Abner Shimony for his connection of Whitehead to quantum state reductions. I am also indebted to Dave Cantrell for artwork, to Michel Weber for suggesting and supporting this work, to Patti Bergin for reference management, and to Christina Heller for manuscript review.

References

Aharonov Y., and Vaidman, L. (1990) Properties of a quantum system during the time interval between two measurements. Phys. Rev. A. 41:11.

Aspect, A., Grangier, P., and Roger, G. (1982) Experimental realization of Einstein-Podolsky-Rosen-Bohm *Gedankenexperiment*: a new violation of Bell's inequalities. Phys. Rev. Lett. 48:91-94.

Bierman DJ, Scholte HS (2002) Anomalous anticipatory brain activation preceding exposure of emotional and neutral pictures http://a1162.fmg.uva.nl/~djb/publications/2002/frmi_presentiment_pa2002.pdf

Cajal SR (1899) Textura del sistema nervioso del hombre y de los verterbrados, See also Cajal, (1997) Texture of the nervous system of man and the vertebrates: Annotated and edited translation of the original 1899 Spanish text by Pedro Pasik and Tauba Pasik. Springer-Verlag, Wien/New York

Chalmers, D. J., (1996) The conscious mind In search of a fundamental theory. Oxford University Press, New York.

Everett H (1973) The theory of the universal wave function. In The many-wprlds interpretation of quantum mechanics, ed. BS DeWitt and N Graham, Princeton University Press, Princeton, New Jersey

Hagan S, Hameroff S, Tuszynski J (2002) Quantum computation in brain microtubules: Decoherence and biological feasibility Physical Reviews E, in press

Hameroff S.R., and Watt R.C. (1982) Information processing in microtubules. Journal of Theoretical Biology 98:549-561

Hameroff, S.R. (1998) "Funda-Mentality" - Is the conscious mind subtly connected to a basic level of the universe? Trends in Cognitive Science 2(4):119-127 (1998).

Hameroff, S.R., and Penrose, R., (1996a) Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. In: Toward a Science of Consciousness The First Tucson Discussions and Debates, S.R. Hameroff, A. Kaszniak and A.C. Scott (eds.), MIT Press, Cambridge, MA. Also published in Mathematics and Computers in Simulation 40:453480.

Hameroff, S.R., and Penrose, R. (1996b) Conscious events as orchestrated spacetime selections. Journal of Consciousness Studies 3(1):36-53.

Hameroff SR (1999) The neuron doctrine is an insult to neurons. Behavioral and Brain Sciences 22(5):838-839.

Hirokawa, N. (1991) Molecular architecture and dynamics of the neuronal cytoskeleton. In pp 5-74. The Neuronal Cytoskeleton, RD Burgoyne (ed.), Wiley-Liss, New York.

Leibniz, G.W. (1768) Opera Omnia. 6 volumes, Louis Dutens, ed. Geneva

Libet, B., Wright, E.W. Jr., Feinstein, B., and Pearl, D.K. (1979) Subjective referral of the timing for a conscious sensory experience. Brain 102:193224.

Marshall, I.N. (1989) Consciousness and BoseEinstein condensates. New Ideas in Psychology 7:73 83.

McGinn, C. (1991) The problem of consciousness, Oxford: Basil Blackwell

Moravec H.P. (1987) Mind Children. University Press, San Francisco

Nagel, T. (1974) What is it like to be a bat? Philosophical Review 83:435-450.

Penrose, R. (1971) in Quantum Theory and Beyond. ed E.A. Bastin, Cambridge University Press, Cambridge, U.K.

Penrose, R. (1989) The Emperor's New Mind, Oxford Press, Oxford, U.K.

Penrose, R. (1994) Shadows of the Mind, Oxford Press, Oxford, U.K.

Penrose, R., Hameroff, S.R. (1995) What gaps? Reply to Grush and Churchland. Journal of Consciousness Studies 2(2):99112.

Rasmussen, S., Karampurwala, H., Vaidyanath, R., Jensen, K.S., and Hameroff, S. (1990) Computational connectionism within neurons: A model of cytoskeletal automata subserving neural networks. Physica D 42:428-449.

Rensch, B. (1960) Evolution Above the Species Level. Columbia university Press, New York.

Rovelli C, Smolin L (1995a) Discreteness of area and volume in quantum gravity. Nuclear Physics B 442:593619

Rovelli C, Smolin L (1995b) Spin networks in quantum gravity. Physical Review D 52(10)5743-5759

Russell, B. (1954) The Analysis of Matter. Dover, New York.

Samsonovich A, Scott AC, Hameroff SR (1992) Acousto-conformational transitions in cytoskeletal microtubules: Implications for information processing. Nanobiology 1:457-468

Scott AC (1995) Stairway to the mind. New York: Springer-Verlag (Copernicus).

Sherrington CS (1953) Man on his Nature, Second Edition. Cambridge University Press, Cambridge U.K.

Shimony, A., (1993) Search for a Naturalistic World View Volume II. Natural Science and Metaphysics. Cambridge University Press, Cambridge, U.K.

Shimony, A. (1997) On mentality, quantum mechanics and the actualization of potentialities, pp. 144-160, in Penrose R (with A Shimony, N Cartwright, S Hawking) The Large, the Small and the Human Mind. Cambridge University Press, Cambridge, U.K.

Smolin, L. (1997) Life of the Cosmos, Oxford Press, N.Y.

Spinoza, B. (1677) Ethica in Opera quotque reperta sunt. 3rd edition, eds. J. van Vloten and J.P.N. Land (Netherlands: Den Haag)

Wheeler, J.A. (1990) Information, physics, quantum: The search for links. In (W. Zurek, ed.) Complexity, Entropy, and the Physics of Information. AddisonWesley.

Whitehead, A.N., (1929) Process of Reality. Macmillan, London.

Whitehead, A.N. (1933) Adventure of Ideas, Macmillan, London.

Woolf NJ, Hameroff S (2001) A quantum approach to visual consciousness, Trends in Cognitive Sciences 5(11):472-447