Is the Brain a Quantum Computer?

Abninder Litta, Chris Eliasmith, Frederick W. Kroon, Steven Weinstein, Paul Thagard

School of Computer Science, University of Waterloo
Department of Philosophy, University of Waterloo
Department of Systems Design Engineering, University of Waterloo
Department of Physics, University of Waterloo
Department of Psychology, University of Waterloo

Received 16 February 2005; received in revised form 8 July 2005; accepted 18 October 2005

Abstract

We argue that computation via quantum mechanical processes is irrelevant to explaining how brains produce thought, contrary to the ongoing speculations of many theorists. First, quantum effects do not have the temporal properties required for neural information processing. Second, there are substantial physical obstacles to any organic instantiation of quantum computation. Third, there is no psychological evidence that such mental phenomena as consciousness and mathematical thinking require explanation via quantum theory. We conclude that understanding brain function is unlikely to require quantum computation or similar mechanisms.

Keywords: Brain function; Consciousness; Explanation; Neuroscience; Quantum computation

1. Introduction

Scientific attempts to understand human thinking have historically drawn on analogies with contemporary technologies, from clockworks to telephone switchboards to digital computers. Today, one of the most exciting emerging technologies is quantum computation, which attempts to overcome limitations of classical computers by employing phenomena unique to quantum-level events, such as nonlocal entanglement and superposition. It is therefore not surprising that many researchers have conjectured that quantum effects in the brain are crucial for explaining psychological phenomena, including consciousness (e.g., Alfinito & Vitiello, 2000; Chrisley, 1997; Hameroff, 1998b; Kak, 1995; Penrose, 1994, 1997).

We argue, however, that explaining brain function by appeal to quantum mechanics is akin to explaining bird flight by appeal to atomic bonding characteristics. The structures of all bird...
wings do involve atomic bonding properties that are correlated with the kinds of materials in
bird wings: most wing feathers are made of keratin, which has specific bonding properties.
Nevertheless, everything we might want to explain about wing function can be stated inde-
pendently of this atomic structure. Geometry, stiffness, and strength are much more relevant to
the explanatory target of flight, even though atomic bonding properties may give rise to spe-
cific geometric and tensile properties. Explaining how birds fly simply does not require speci-
fying how atoms bond in feathers.

The primary aim of the cognitive sciences is to provide explanations of important mental
functions, including perception, memory, language, inference, and learning. We contend that
quantum properties are irrelevant to explaining brain functions, just as bonding properties are
irrelevant to explaining wing function. Compelling explanations describe mechanisms that
generate the phenomena associated with the explanandum. Mechanisms are systems of entities
and activities organized such that they produce regular changes (Machamer, Darden, & Craver,
2000). Neural mechanisms in particular consist of neurons, neuronal groups, and functional
brain areas whose biological–computational activities generate brain operations. A neural
mechanism is a plausible explanation of a mental phenomenon if there is biological and psy-
chological evidence that the entities and activities that it identifies are in fact responsible for
the changes that constitute the phenomenon. We argue that neurocomputational rather than
quantum mechanisms provide the most credible explanations of mental phenomena.

We present three classes of reason why it is implausible that quantum mechanical processes
are relevant to explaining how brains operate. Our first argument is computational: that quan-
tum mechanisms are unlikely to play a role in information processing in the brain. Our second
argument is biological: that there are several reasons why the essential functionality of an or-
ganic system such as the brain should not require quantum mechanical explanation. Finally,
our third argument is psychological: that there is no reason to believe that quantum computing
contributes to mental phenomena. We conclude that the conjecture that the brain is a kind of
quantum computer is inferior to the neurocomputational hypothesis, which states that the brain
produces mental phenomena by means of neural processes that encode, transform, and decode
information represented by patterns of neural activity (e.g., P. Churchland, 1989; P. S.
Churchland & Sejnowski, 1992; Eliaismith & Anderson, 2003; Rieke, Warland, de Ruyter van
Steveninck, & Bialek, 1997; Smolensky, 1994; Thagard, 2005).

2. The computational argument

The theory of quantum computation originated with the pioneering ideas of Feynman
(1982) and the universal machine proposed by Deutsch (1985). Quantum computing employs
qubits, whose states, unlike those of bits in classical digital computers, may exist simulta-
neously as coherent superpositions of both 0 and 1. Potentially, quantum computers may em-
ploy a unique sort of parallelism that would make tractable some problems currently consid-
ered intractable via classical computation. In computational complexity theory, a problem is
considered tractable if it can be solved in polynomial time, that is, if the time required to solve
it increases asymptotically by at most a polynomial function of the size of the input. Problems
with optimal solution times increasing faster than this (e.g., as an exponential function of the input size for sufficiently large values) are considered to be intractable.

The technological potential for quantum computing was first realized in the formulation by Shor (1994) of a polynomial-time quantum algorithm for the problem of factoring a number into its constituent primes, for which the best classical algorithms require exponential (or at least superpolynomial) time. The apparent intractability of prime factorization under classical computation is central to the security of the cryptographic schemes prevalent today. If implemented on a suitable quantum machine, Shor’s algorithm could potentially break the encryptions depended on by governments, banks, and millions of individuals.

Indisputably, phenomena requiring quantum mechanical explanation exist throughout the brain, and are fundamental to any complete understanding of its structure and physical mechanics. Every molecular bond and chemical interaction is ultimately explained by quantum theory. Other examples of the relevance of quantum effects to biological systems include explanations of enzyme energetics (Welch, 1986), cross-membrane electron transport in photosynthesis (Vos, Rappaport, Lambry, Breton, & Martin, 1993), and, in the brain itself, the threshold-breaching effects of a single neuronal ion channel opening (Johnson, 2001). However, none of these effects contribute essentially to explaining the overall functionality of the associated system, which can be fully described without explicit appeal to quantum-level phenomena. In our wing analogy, it is unnecessary to refer to atomic bonding properties to explain flight. We contend that information processing in the brain can similarly be described without reference to quantum theory. Mechanisms for brain function need not appeal to quantum theory for a full account of the higher level explanatory targets.

Compare the ordinary digital computer, in which quantum effects are central to understanding electron flow through transistors and low-level circuitry. Indeed, one of the motivations for pursuing quantum computing is that chip designers are rapidly approaching quantum limits to how small circuits can be printed. Nonetheless, such quantum phenomena play no role in the computation performed by the system. The base unit of information processing in the computer is the logic gate. Gates are implemented in transistors, which are designed to be noise tolerant and to correct for anomalies caused by individual quantum events. Digital computers are thus designed to ensure that quantum effects do not contribute to information processing, whereas in quantum-based machines such properties as superposition and entanglement play a fundamental role in computation. The system of chips, transistors, and logic gates thus constitutes the appropriate mechanism for explaining the principal functions and behaviors of digital computers, and quantum-scale mechanisms are irrelevant.

Similar considerations hold in the brain. There is substantial evidence that the timescales for individual quantum events in the brain are not in accord with the temporal requirements for influencing neural firing in a consequential (i.e., non-noise) manner. Eliasmith (2001) reviewed empirical evidence that demonstrates that specifying firing times within about 1 msec captures nearly all of the available information in a neural spike train. This is because unreliable presynaptic vesicle release, disparate neurotransmitter concentrations in these vesicles, and neural spike timing “jitter” induced by axons limit the possible precision of neural information transfer. As a result, environmental noise severely constrains the useful temporal precision of neural spiking in most of the cortex. The fastest neural timescale sensitivities are found in highly specialized subsystems, such as the barn owl auditory system. Even there, however, pre-
cision is only on the order of microseconds (Carr & Friedman, 1999). Calculations by Tegmark (2000) explained why this implies that effects on the timescale of quantum phenomena can be safely averaged out as noise.

Specifically, Tegmark (2000) estimated that, at normal brain temperature (~310 K), individual neurons have decoherence timescales of $10^{-20}$ sec. For a neural microtubule, a cellular substructure that might be a more plausible component of quantum computation in the brain, Tegmark obtains a $10^{-13}$ sec timescale. Both derivations ignore factors such as water molecule collisions and the effects of other ions, such as Cl—, which would further shorten these times. In contrast, the fastest firing neurons work on millisecond timescales, and polarization excitations in even the shortest microtubules are of order $10^{-7}$ sec. Thus quantum-level events, in particular the superpositional coherences necessary for quantum computation, simply do not have the temporal endurance to control neural-based information processing. Indeed, if neurons and their associated spike patterns are at all involved in consciousness or any other cognitive process, then such processes cannot be quantum computational. The large timescale discrepancies described by Eliasmith (2001) and Tegmark (2000) rule out any substantial impact of quantum-level phenomena on neural firing patterns beyond the level of correctible noise. Although it could perhaps be argued that extremely short quantum events somehow “restructure” neurons or neural interactions, to effect changes at the timescale of spiking, these speculations are hampered by the significant biological plausibility problems we explore in the next section.

Hence, explaining information processing in the brain via neural mechanisms does not require an account of individual quantum-level events taking place in the brain. Of course, we are not claiming that brain function is somehow divorced from the quantum mechanical bedrock of physical reality. Rather, such effects simply do not have explanatory relevance to the question of mental operation, as with our bird-flight analogy. Neurocomputation is relatively robust with regard to such phenomena, in that input–output relations are unaffected by slight variations of them. Certainly, explanation of why a particular neuron fired at a precise moment may require reference to quantum mechanical phenomena that affected a specific ion channel. However, just as a digital computer can be described independently of the fact that the number of electrons in a single logic gate at a specific time is fundamentally quantum, a functional explanation of the brain need not resort to quantum mechanisms. Although the functional role of neurons is somewhat different from that of the transistors and logic gates of a digital computer, for the general operations of the brain, quantum effects are at a low enough level that any associated fluctuations can be categorized and handled as noise.

### 3. The biological argument

Significant progress is being made in the design and production of large-scale quantum computers, and operational machines employing up to seven qubits have already been constructed (Vandersypen et al., 2001). Nevertheless, such machines require working conditions that contrast vividly with the immediate environment of the brain. The power of a quantum computer lies in its ability to maintain superposed qubit states long enough to facilitate superparallel computation. The maintenance of an extraordinarily high degree of isolation
from even minute environmental interactions is a vital prerequisite for preventing decoherence, which is the decay of coherent quantum-state superpositions caused by such interaction. Exceedingly low operational temperatures are also a necessity for most physical implementations of quantum computers, although simpler machines based on nuclear magnetic resonance have managed room-temperature coherence over useful timescales (Cory, Fahmy, & Havel, 1997).

Standing in stark contrast to these physical requirements are the conditions that exist in and around animal brains. Brains are warm, wet, biological constructs, honed by evolution to exhibit the sort of robustness and durability needed for survival in the world. Although to some extent they are protected from the environment by a thick skull, cushioning fluid layer, and so on, this isolation is nowhere near sufficient to maintain large-scale quantum coherence at the neuronal level for computationally significant periods of time (Nielsen & Chuang, 2000; Tegmark, 2000). Interactions with endogenous secretions and structures, changes in external temperature, or even routine physical trauma such as moderate blows to the head would cause more than enough disruption to the internal brain environment to end any sort of nontrivial quantum coherence.

Some have argued that tiny protein structures within neurons, microtubules, offer a milieu suitably sized and isolated for quantum coherence and computation (e.g., Hameroff, 1998b; Kak, 1999; Nanopoulos, 1995). But these theories lack any empirical support and also run afoul of the previously mentioned decoherence–neural spike timescale discrepancies. Moreover, they raise the question of what makes brain microtubules so special that they alone allow for quantum computation. Microtubules are generic cellular structures that are involved in internal transport, combine to form cilia and flagella, and play a proven role in maintaining cytoskeletal structure (Grush & P. S. Churchland, 1995). Found throughout the plant and animal kingdoms, their distribution in neurons is wholly unexceptional. Indeed, the hypotheses regarding microtubules offer nothing equivalent to traditional neuroscientific explanations of interspecies disparity. For instance, Allman and colleagues (Allman, Hakeem, Erwin, Nimchinsky, & Hof, 2001; Allman, Hakeem, & Watson, 2002; Nimchinsky et al., 1999) identified spindle neurons, found only in humans and the great apes. These unusually large cells seem to act as emotional and motivational relays and show high activation in demanding tasks involving self-control and recognition of one’s errors (Allman et al., 2002). Humans and our closest relatives can perform these tasks better than other animals, and spindle neuron concentration among hominid species seems strongly correlated with such success as well. In contrast to this well-supported neuronal explanation, quantum–microtubule theorists have yet to outline plausible mechanisms by which species differ in their abilities. In this absence, are we to believe that carrots and rutabagas also exhibit quantum computation, or are conscious? As P. S. Churchland (1998) argued, “The want of directly relevant data is frustrating enough, but the explanatory vacuum is catastrophic. Pixie dust in the synapses is about as explanatorily powerful as quantum coherence in the microtubules” (p. 121).

Another significant physical obstacle to quantum computation in the brain is the matter of error correction, which pertains to noise tolerance in the transmission and processing of information. Error correction in the brain is a real phenomenon, and several possible neural correlates have been proposed with empirical support (e.g., Smith & Shadmehr, 2000; Stiber, 2005). Although redundant networks may also play a role, the most common brain implementations
of error correction and recovery seem to involve either tuned attractor boundaries or high-precision spike codes—both well-understood engineering concepts (Stiber & Holderman, 2004). The task is similarly straightforward under a classical redundancy-based computational scheme. For example, digital computers employ redundant, multiple-bit encodings to allow for the correction of a minority of bit values erroneously flipped by noise sources. This simple approach is unworkable in a quantum computer, because the perfect replication of an unknown qubit state is impossible (Dieks, 1982; Wootters & Zurek, 1982). Moreover, one cannot even examine qubits to see if corrections are needed without causing measurement-induced superposition collapse. The common solution outlined in Nielsen and Chaung (2000) involves a complicated scheme of entangled qubits and joint measurements that reveals errors without spoiling computation. But it is implausible that such a contrived strategy developed naturally in biological systems, given the strong physical and environmental obstacles to bioinstantiation. Even Kak (1999), a proponent of the idea that the brain is a quantum computer, admitted that the currently proposed error correction techniques “work under very artificial and unrealistic assumptions” (p. 10). Although we acknowledge nature’s ability to evolve ingenious solutions to difficult problems, the burden of proof is on those who would invoke quantum mechanics to not only provide the details of such a biological mechanism for quantum error correction, but to do so in the face of physical evidence for simpler, classically based alternatives.

Even if quantum computation in the brain were technically feasible, there is a question about the need for such massive computational efficiency in explaining the mind. It is technologically desirable that a quantum computer should factor a large number into primes in polynomial time, but there is no evidence that brains can accomplish this or any similarly complex task as quickly. As well, the fact that certain quantum algorithms can be asymptotically better than classical counterparts says little about how the algorithms compare on typical input sizes, which could in fact be quite small. For instance, given problem input size $n$, a polynomial solution requiring time $9000 \times n^7$ would be slower than an exponential solution requiring time $2^n$ for $1 < n \leq 53$, a range which may encompass all conventional input sizes. In everyday applications, then, quantum does not necessarily mean faster. Although it could be advantageous to have massively parallel quantum brains, it seems very unlikely that natural selection would evolve a vastly complicated and fragile system of information processing for no compelling survival purpose. No evidence has been generated that brains need the power of quantum parallelism to support the basic biological needs of survival and reproduction. Even if it might somehow be useful, the substantial environmental obstacles described previously make the natural evolution of organic quantum computation an implausible notion. Moreover, as we argue next, there is no compelling evidence that we need quantum computing or related processes to explain any observed psychological behavior, such as consciousness.

4. The psychological argument

Although the hypothesis that the brain is a quantum computer is biologically and computationally implausible, there might be psychological phenomena not amenable to a neurocomputational explanation that are explicable by appeal to quantum theory. Penrose (1994,
1997) and Hameroff (1998a, 1998b) argued that mathematical thinking and conscious experience are two such phenomena.

The Penrose–Hameroff “Orch OR” model for brain function describes consciousness as depending fundamentally on the orchestrated, noncomputable wave-function collapse (reduction) of coherent quantum states in neural microtubules (Hameroff, 1998b). The collapse must not be completely random, as would be the case with environmental decoherence, so Penrose (1997) postulated the existence of an objective reduction (OR) phenomenon based on “self-collapse” due to quantum gravitational effects on space time. Before the coordinated collapse that signifies conscious processing, quantum-coherent states within different microtubules are thought to interact with one another, thus achieving a sort of rudimentary quantum computation in the brain for preconscious processes.

There are many problems with this theory, beginning with the OR conjecture. Penrose’s (1994, 1997) idea that superpositional collapse happens on its own, independent of environmental interaction, is an extremely controversial one. Hawking (1997) voiced opposition to the idea, arguing “that warping [of space-time under OR] will not prevent a Hamiltonian evolution with no decoherence” (p. 170). More general resistance centers on the implications of the hypothesis. The existence of OR in the manner described by Penrose and Hameroff (1998b) would require fundamental and far-reaching revisions to quantum theory itself. Although Penrose freely accepted this fact, such a significant claim requires substantial supporting evidence. Even discounting the lack of direct empirical support, however, the main theoretical arguments for the necessity of OR are also critically flawed.

First, Hameroff (1998a) proposed an account of anesthetic action as evidence for a quantum mechanical theory of consciousness. He argued that anesthetics bind in hydrophobic pockets of certain brain proteins, preventing quantum delocalization and macroscopic coherence (and thus consciousness) by inhibiting the quantum states created by the London forces between amino acid groups that are fundamental to the conscious (nonanesthetized) brain. Until recently, no alternative, purely biochemical theories of general anesthesia were available.

In the past few years, however, there has been an explosion of experimentally supported explanations of general anesthetics based on molecular biology (Moody & Skolnick, 2001). The previous hypothesis of hydrophobic action, that anesthetics interact with cell membrane lipid, has been abandoned in favor of mechanisms involving protein ion channels in the brain, particularly ligand-gated channels such as receptors for GABA, NMDA receptors, and nicotinic acetylcholine receptors (Flood, 2002). These hypotheses have substantial empirical support (e.g., Garrett & Gan, 1998; Siegwart, Krähenbühl, Lambert, & Rudolph, 2003; Williams & Akabas, 2002). As Flood stated, “Every general anesthetic in use today acts on at least one type and in some cases several types of ligand gated ion channels” (p. 153). None of the proposed explanatory mechanisms involve quantum mechanical properties or quantum computation. Anesthesia thus provides no empirical support for the speculations of the Orch OR model with regard to consciousness.

Orch OR might nevertheless be needed if there were no competing explanations of the phenomena associated with consciousness. However, just as Hameroff’s (1998a) quantum mechanical theory of anesthesia has been surpassed by biochemical explanations at the molecular level, we expect that quantum theories of consciousness will be superseded by the development of neurocomputational explanations. Already, there are neural-based hypotheses that
provide steps toward an understanding of conscious experience (e.g., Crick & Koch, 1990, 1998; Damasio, 1999; Steriade, McCormick, & Sejnowski, 1993). The scientific exploration of consciousness is still in its infancy, but there is no evidence to suggest any superiority of quantum mechanical over neurocomputational explanations. Orch OR in particular, considered the most detailed quantum brain theory, offers neither brain-based empirical support nor any compelling explanatory mechanisms for mental phenomena (Grush & P. S. Churchland, 1995).

Our critique of the plausibility of quantum–computational explanations of consciousness does not purport to prove that such explanations are impossible. Further theoretical and experimental research may indeed develop principles and evidence that tie aspects of consciousness, such as shifts of attention or qualitative experience, much more closely to quantum-level processes. However, given the computational and biological problems discussed in the previous two sections, we think that such developments are less likely than the continued exploration of standard neurocomputational explanations that are better supported by experimental data. Our implausibility arguments stand against any information processing in the brain utilizing uniquely quantum mechanisms. If consciousness plays a role in information processing, our critique applies to quantum theories of consciousness as well. Our arguments may thus be unacceptable to those who take consciousness to be nonphysical, purely psychological, or outside the realm of scientific explanation, but this debate over materialism is beyond the scope of our analysis of the physical aspects of quantum descriptions of the mind.

The other psychological phenomenon that has been prominently advocated as requiring alternative explanation is mathematical thinking. Penrose (1994, 1997) argued on the basis of Gödel’s first incompleteness theorem, which shows the incompleteness of any consistent formal system for arithmetic that mathematical insight is fundamentally noncomputable, and therefore requires the OR phenomenon and associated quantum computational processing in the brain. Numerous respondents have demonstrated, however, that Gödel’s theorem does not have the implications drawn by Penrose (e.g., Grush & P. S. Churchland, 1995; LaForte, Hayes, & Ford, 1998; Manaster-Ramer, Zadrozny, & Savitch, 1990; Shapiro, 2003). Although we are still far from having a neurocomputational theory of mathematical reasoning, Gödel’s theorem does not imply that mathematical insight must be noncomputable. Cognitive science has generated some interesting hypotheses about how minds generate and appreciate mathematics (e.g., Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Gallistel & Gelman, 2005; Lakoff & Núñez, 2000), but no reasons have been provided for quantum computation to be in any way involved in the psychology of mathematics. Simply theorizing about possible quantum effects in the brain is not enough: Abandoning the promising neurocomputational framework in favor of Orch OR requires the existence of compelling mental phenomena that demand alternative explanation. Moreover, the Orch OR model would have to provide detailed and biologically sound mechanisms for these alternative explanations; it is at present more of a promissory note than a proper model.

In sum, we have provided an interlocking set of computational, biological, and psychological arguments against the hypothesis that the brain is a quantum computer. Let us return once more to our bird-flight analogy. The relevance of atomic bonding properties to the structure of wings does not necessitate their involvement in explaining flight, because aerodynamic mechanisms have proven sufficiently powerful to explain the phenomenon. Only if specific,
flight-relevant geometric or tensile features arose purely from atomic bonding properties in feathers would it make sense to import these details into our explanations of bird flight. Because no such special properties are found in existing examples of wings, atomic bonding is not relevant to explaining bird flight. Similarly, there appear to be no special quantum mechanical properties needed to explain psychological and neurological phenomena. The onus is on those who would appeal to quantum theory to show the existence of aspects of the brain that are not explained by neurocomputational theories, and that can be explained by quantum computation or associated mechanisms. Although the discovery of solid evidence for fundamentally quantum characteristics of mental phenomena would be tremendously exciting, current ideas fall well short of this standard.

Notes

1. Hameroff argued that the model he has developed with Roger Penrose is immune to Tegmark’s critique, because their theory is based on superposition at the level of the constituent proteins of microtubules, which have longer possible coherence times (Hagan, Hameroff, & Tuszyński, 2002). Yet forcing this descent to below even the level of microtubules further complicates the already implausible mechanisms of the Penrose–Hameroff model. See Section 4 for further details.

Acknowledgments

We thank Darren Abramson, James Ladyman, Kelly John Rose and Durgesh Saraph for comments on earlier versions. This research was supported by the Natural Sciences and Engineering Research Council of Canada.

References


