THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2012-09

Consciousness, Time and Complexity

Vahid R. Ramezani

The Institute for Systems Research

A. JAMES CLARK SCHOOL OF ENGINEERING

ISR develops, applies and teaches advanced methodologies of design and analysis to solve complex, hierarchical, heterogeneous and dynamic problems of engineering technology and systems for industry and government.

ISR is a permanent institute of the University of Maryland, within the A. James Clark School of Engineering. It is a graduated National Science Foundation Engineering Research Center.

www.isr.umd.edu

Consciousness, Time and Complexity

Vahid R. Ramezani

Institute for Systems Research,

University of Maryland, College Park, MD 20742, rvahid@umd.edu

August 20011; revised August 2012

Abstract

We explore the connection between the mind and the brain. We propose that consciousness is the consequence of processing information and that the solution to the binding problem does not entail quantum mechanical coherence or entanglement. We argue for an alternative inspiration from quantum mechanics and quantum field theory based on time-energy uncertainly: not to reduce consciousness to a quantum wave function but to see what quantum mechanics teaches us about information, time, complexity and transformation. We introduce three postulates and a law governing cognitive systems.

1 Introduction

It has been proposed that quantum mechanics plays a central role in the essential functions of the brain [1]-[2]. Quantum mind's basic idea is that quantum superposition provides a solution to the binding problem. Max Tegmark has shown quantum coherence is probably too fragile to account for neural activity and the brain is essentially a classical object [3].

The attempt to apply quantum coherence to the binding problem reflects the assumption that the information content of the Universe is fundamentally fragmented. Not only do thermal properties of the brain prevent the application of quantum coherence and quantum entanglement to the binding problem, quantum entanglement itself might be a manifestation of unified properties of the universe and not a mechanism for it, quantum superposition an abstraction, incomplete in its description of reality [4].

Neural correlates of cognitive processes have been extensively studied and our discussion below is only intended to open the way to some philosophical ideas.

The central observation about the brain is that certain properties of it, according to experiments, are fundamental to consciousness. These are chemical and electro-chemical processes at neurons, glial cells, synapses and along the axons.

We propose that the essence of consciousness is information processing in space-time. Consciousness is the experience of transformation and interpretation of information in a highly intergraded and space-time compact brain. “Compactness” is described when we invoke certain optimization principles.

The sense of conscious unification of information being processed in a coherent manner at a classical level reflects a fundamental property of the universe. The character and extent of such unification depends on local properties but its essence does not. The universe is whole.
We now know that living organisms utilize quantum effects and these effects might influence the quality of our cognition; for example, we might one day discover that they play a role in the function of neurotransmitters without being fundamental to the binding problem.

2 Thought experiments

Suppose we manage to create a system composed of two brains intertwined in the same location of space where the information signals propagating throughout one brain is completely isolated (in the sense of information) from the other. Will one mind be aware of the other? Or influence it in any way? Suppose we could scale up the brain so that the overall signaling and information processing would remain invariant but information signals would have to travel much longer paths. Will such a machine have the same quality of awareness as our own? Lastly, suppose we could freeze the activities of a machine which is in every aspect similar to the brain and start them again in time-steps years apart; processing the same information content.

The first thought experiment might not be far from reality. The human brain, like other biological organs produced by evolution, is highly redundant. A damaged brain can sometimes reconfigure itself to regain some of its previous functions. We could imagine, architecturally, several processing networks capable of creating consciousness which are intertwined in our brain.

We know that space-time distance is crucial to interactions between particles. Particle interaction is an example (though, in our view, not the basis) of unified character of the universe. The Hamiltonian acting on the tensor products in quantum mechanics is a unification of particle interactions. While the harmonic oscillator of spring mass system can be broken down into a series of action and reactions at smaller scales, all the way down to atomic and molecular forces, quantum oscillator cannot; it feels fundamental and it probably is. This is in fact the underlying reason one might seek quantum mechanical or quantum field effects. The idea is that perhaps we could write down something like a Hamiltonian or a field propagator for the brain and generate a very compact space-time picture which corresponds to essentially simultaneous integration of information at a sharp cross section of time, solving the binding problem.

3 Information patterns of the brain

Let us describe a plausible philosophical picture and examine the thought experiments:

To experience consciousness, the neurons must fire; some exciting others, some inhibiting others, synchronizing their firing rates and intensities, coupling and decoupling from each other etc. First, consider flow and transformation of a single information pattern through the Neural Correlates of Consciousness (NCC) [5]. A pattern is composed of atomic information actions of the brain (firing of neurons etc.) acting on data representations which are replicated or transformed. The replication is not a simple mechanical process like water waves traveling

through water (though this might be a useful analogy). The reconstruction of a pattern is complex, similar to drawing a figure by hand requiring scanning the entire picture as we draw it with repeated corrections and feedbacks via comparisons to the original. It would entirely be clear to an observer that the pattern is being carefully replicated (such as a photograph being turned into a cartoon where essential features are maintained and emphasized) not propagated or transformed "automatically" via elementary laws of physics.

The pattern is unified by its coherent propagation and interpretation. It enters our subjective experience because of this very propagation and transformation as a whole. This is an information property independent of the specifics of the underlying physics. The conscious experience results from achieving a critical level of complexity and purpose in transformation and propagation of the pattern.

How are different patterns put together? Once again, we think of communication and information binding. We can imagine a constant (possibly highly complex) "background" pattern in NCC. (This might be the pattern we experience when sensory inputs are cut off and we are able to clear our mind from memory. When it is maximally coherent and symmetric, we might lose track of temporal and spatial relations) When two patterns in different parts of NCC are propagating; they "perturb" this uniform background, sending ripples throughout the network with "pilot waves" that can reach and alter all existing wave patterns, perhaps slightly, by adding and encoding some properties of one on the others. We experience this as presence of all patterns in our mind in a unified state. The actual correlation in the real world might not exist. When we look at roses and smell them; it is our mind that creates a correlation. We can imagine that patterns share some network paths; in which case, further communication is required for managing the superposition of patterns.

The actual situation is far more complex of course. For example the NCC might produce tightly coupled multiple copies of a pattern which are then transformed or tested to check for different properties. We might not feel these multiplicities. Since NCC is the very expanse of our consciousness, we have no larger context for examining this question. We might in fact experience these multiplicities and they might precisely be what we think as a single entity.

The precise space-time scales are difficult to ascertain. It follows from the information processing perspective that time must be subjugated to information. The time interval the brain requires to propagate or interpret a pattern with all the events relating to these actions are unified, past is not lost. For example, the very notion of neuron firing rates only makes sense with the coupling of the past with present as time cycles are required to measure rates. However, a transformed pattern replaces its old version in NCC and the past associated with it is lost. Because only propagating patterns are perceived, the past is truncated to the time required for transformation-propagation processes.

If the brain is placed in a very strong gravitational field (and assuming it survives) we would observe that NCC processes last hours instead of fractions of a second. Most likely, the unification will not break down. However, stopping the process would disrupt the transformation and restarting the process later can only define a different transformation. The unification would break down. The precise extent of temporal and spatial unification is difficult to characterize and

might be subject to certain cognitive laws of uncertainty. It might be fundamental to the universe and we might ultimately have to refer to the brain itself tautologically. The limits of the synchronization rate might limit space-time scanning of a pattern causing a pattern transformation to split into a collection of sub-transformations. Let us examine this more carefully.

Think of the brain's information patterns as a space-time cylinder with space being 2-dimensional for simplicity. A pattern is a spiral of actions on data which collapses as it unifies across space-time. Vertically (in time), how long the (uncertainty) in the spiral can grow before it collapses not into one, but more than one, with only some actions sufficiently concentrated in time unifying in one pattern and the rest lost to that pattern? "*When all the earthly things have been destroyed, how much does the soul of the past persist*?"

Ideas and speculations might be inspired by quantum field theory [6] and unified interaction of particles. The behavior of interacting particles obeys the Lagrangian of their respective fields and (generalized) optimization of actions. It is perhaps the single most fundamental idea we have about the character of the universe. Optimization, given certain constraints, is related to symmetry and emergence of structures. It is possible that NCC could have only achieved their transformational complexities and allowed to become unified in space-time by obeying certain optimal actions given the basic design of the brain. The conscious transformational aspects of NCC might not survive arbitrary scaling. A neuron or a group of neurons firing today and expecting to be unified with the ones firing tomorrow might be violating the scales by which information *can be* processed optimally given some constraints. Time is a real relational entity and perhaps consciousness could not have come about without optimal degree of delineation and compactness in time; i.e., allowing the past to expire might be a condition for consciousness. The past might be allowed to persist only as much as uncertainties associated with synchronization of the basic patterns of the brain require.

Is the influence of an earlier note when we listen to music a manifestation of this temporal unification? It follows from our argument that this cannot be the case. The sensation is most likely created by "bouncing" the earlier note from memory and creating a coupled pattern across space with the present note, not across time. The temporal unification is never felt across time; it is temporally minimal and because a note already represents temporal unification of time cycles of sound waves, the pattern has already collapsed because we can hear a single note in an alert state. In fact, even a single note is felt across time first. As we hear it, we spread it across space, possibly using our memory, and collapse it again, in minimal collapse time, creating a more complex pattern and sensation.

This leads us to an inevitable conclusion: If the mind can delineate patterns in time, then they cannot be temporally collapsed into one entity, unless they are superposed, coupled and spread across space perhaps using memory. Thus, the fundamental temporal collapse time-interval is smaller or equal to our mind's finest *cognitive* perception of time.

We return to our first thought experiment of two intertwined brains. From our reasoning, it follows that if the information synchronizing waves break down between the two brains, two independent minds would emerge even if physical signals such as thermal fluctuations still

connect the two. The NCC's information signals can carry information and generate meaningful correlations between patterns. It is not simple spatial proximity or thermal equilibrium but information processing that binds patterns.

4 Tangible outcomes and free will

We have simply assumed that unification associated with subjective experience is fundamental to the processing of information at classical scales and the mind is not reducible to particle or field interactions. We have done so clearly at an axiomatic level.

A hard problem for us is the characterization of the consequences of the subjective experience, its tangible manifestation. We all know it has consequences. The problem is that if we assert that no instantaneous field and particle state exists which captures the whole of our subjective experience, then how is consciousness revealed to the outside world? The tangible universe is composed of fields and particles; for our awareness to be more than an intangible property it must act on matter. Such actions are highly local in space and time, internal signals to the memory or output signals to the body.

First, we assert that consciousness is far less consequential tangibly at any cross section of time than it might at first appear. Suppose we ask someone: if you are self-aware raise your hand! The person grasps the deep meaning of the question and raises her hand. We accept that the action does reflect the state of being aware of oneself but does it encode its essence or anything significant about it? The action simply demonstrates that the information processes that gave rise to consciousness are also capable of generating a signal after certain processes were completed: hearing the question, comparing it to the memory, integrating various patterns. Our mind is simply capable of producing a state at a cross section of time that reveals some drastic reduction of its intangible properties. When we convey more sophisticated information, we rely heavily on our memories. The processes of integration and interpretation at the highest levels of our cognition have evolved to produce certain outputs. They also produce the subjective experience; one is concentrated in space and time; the other is not. The output is one possible result of these processes and since these processes and our subjective experience resulting from them are equivalent, the output appears and indeed is the result of the subjective experience itself. As an organism or machine reveals such properties, we might conclude it is conscious. This is how we come to believe other people are conscious. More sophisticated forms of self-expression integrate more complex information from memory and sensory inputs. But the conscious experience can only add small amount of information to memory at any cross section of time.

The outcome and the outputs of NCC processes are patterns and signals localized in space-time. The uncertainties of outcomes implied by quantum mechanics and the non-computability of complex processes mean that, at least sometimes, our actions cannot be predicted by any algorithm. In this sense, we have free will. A realist interpretation of quantum mechanics and a deterministic but non-computable picture of reality do not restrict our minds to predictable outcomes. Not even our mind can predict all of our actions.

5 From AI to consciousness

The view that our subjective experience of consciousness is an information property and does not rely on specific physics of the brain implies that machines can achieve it. It is likely that some aspects of the brain architecture and functions are essential. They include notions of information as transformation (a duality of data and process), spatial and temporal correlation, space-time information "collapse", optimization, synchronization, memory, reward potentials etc. However, there might not exist, even in principle, a theory of consciousness or an algorithm for putting all the subtleties together.

Fortunately, we have working examples. We know that primitive organisms had simple sensory structures which grew more sophisticated over time because interpretation and interaction with memory gave them an evolutionary advantage. The ability to look at the outside world was transformed to the capacity to look inward as the neural layers became more complex; data and process were coupled in an upward evolutionary duality. The sense of self-preservation led to the sense of self-awareness. The brain's highly parallel, highly connected architecture can one day be emulated but, it might be that only machines with a primitive sense of self-preservation and competition can eventually become self-aware.

Appendix: Toward Quantification

Consciousness corresponds to transformation of information in time. From this basic observation, we argue for an alternative inspiration from quantum mechanics and quantum field theory: not to reduce consciousness to a quantum wave function but to see what quantum mechanics teaches us about information and transformation. We have implicitly assumed that the information pattern associated with stream of consciousness is unique and "quantized" in time. There are experiments and studies that appear to support this perspective [10] but our approach is axiomatic.

We stated earlier that quantum mechanics and field theory appear to be fundamental; their further refinement being non-computable. Indeed, recent work [7] suggests quantum theory is close to optimal in terms of predictive power.

Our starting point is the well known time-energy uncertainty,

$$\Delta T \, \Delta E \geq h / 4\pi.$$

Time operator does not exist in quantum mechanics and ΔT, the spread in time, is usually interpreted as the "smallest" time interval required for a measurable change to appear in the evolution of a quantum system using some measurement operator A.

$$\Delta T = \frac{\Delta A}{|\frac{d}{dt} < A >|}$$

For us the important observation is that ΔE (the spread in energy) controls the evolution speed in time. Time-energy uncertainty means that a more "complex" system (with more energy eigen states) exhibits smaller minimal times for its evolution. Generally, entanglement, which can be viewed as "unified composition" through tensor products, speeds up evolution of the composed quantum state in line with the time energy uncertainty [9]. Intuitively, tensor product "connects" all degrees of freedom or dimensions from one system to another. We can visualize this with Kronecker product of two matrices where each element of the first matrix is replaced by the second matrix, scaled by that element. Every dimension from the first "interacts" with every dimension from the second and hence the dimension of the product is the product of the dimensions. Tensor products are related to "full connectivity". We can see this more clearly in the graphical definition of tensor products: vertex set of the tensor product G × H is the Cartesian product of vertex sets of the graphs V(G) × V(H); and any two vertices (u, u') and (v, v') are adjacent in G × H if and only if u' is adjacent with v' and u is adjacent with v. So while the nodes of the two graphs are combined in every possible combination, the composition does not change the connectivity structure of the components. We will see a natural and suggestive connection between tensor products of graphs and a cognitive law governing composition.

Brain activity is complex. Although we do not know exactly what kind of complexity corresponds to consciousness, we assume its degree is quantifiable. We visualize this complexity as a four dimensional space-time pattern, capturing spatial-temporal symmetries and correlations. Although a proper description might require vector bundles and tensors (because a point in space, corresponding to a neuron for example, depicts a complex object with smoothly varying properties) combined with fractal geometry (because of unification-self similarity properties to be described shortly).

Consider a collection of oscillators in space coupled by short range and long range transfer of information; i.e., their influence graph is not localized and restricted to neighboring nodes. The oscillators need not have uniform properties and the system might include aggregate transfer functions from one region to another according to some spatial partition. In addition, the whole system could be subject to certain constraints. The partition and influence functions themselves could evolve in time. Many systems of this type, even with only localized influence functions and few components, exhibit chaotic behavior. Perhaps such a system captures some properties of the mind. There is however a key difference: as we slide a time window of finite duration over the space-time patterns generated by such a mechanical system, we would find no particular position of the window to be special. This will not be case with brain activity. The dynamics of the brain are *interpretive* and thus are highly heterogeneous. A window that contains different incomplete pieces of various interpretations is fundamentally different from one which contains them whole. The image of a continuous stream of consciousness flowing like a river requires a meaningful smoothly sliding time window over brain's space-time dynamics. Because the mind has (memory enabled) interpretive mechanisms for continuous movement and the mind has no

way of perceiving its own discrete nature, we have the illusion of continuity in time and a continuous stream of consciousness.

We assume the *information complexity* of the mind is a quantifiable concept and state our postulates:

1) Consciousness is quantized in time according to a fundamental collapse time.

2) The activity of a (conscious) cognitive system must occur at a minimum level of space-time density - a minimum *information complexity density* – and must achieve a minimum *total information complexity*.

3) A more *powerful* (conscious) cognitive system, one with higher total information complexity, cannot have a larger collapse time when compared to a weaker one. (note: achieving higher volume of cognitive information requires higher density if collapse time decreases).

Vertebrates and a few invertebrates might be conscious at some levels (exactly which ones and what levels are not clear but it stretches the imagination to believe that great apes, elephants and dolphins are not conscious) and humans likely sit atop the hierarchy of cognitive powers. Let us assume arbitrarily that fish have the weakest possible degree of consciousness.

We have argued for temporal quantization of the mind. The following thought experiment might strengthen our argument. Suppose a copy of your brain suddenly comes to existence at time zero. Suppose at the cross-section of time when consciousness is achieved first, your mind sends a signal to itself, directly or through memory paths, conveying the belief: "what I think exists at present really belongs to the past". If consciousness was an accumulative stream then you would know (not through some abstract idea but through direct conscious perception) that what you see and perceive "now" as an integrated whole really belongs to the past. In principle, you would be able to delineate past from the future by a conscious mechanism while perceiving it as the "present", a cognitive contradiction.

Our subjective perception of time depends on temporal collapse. A more complex brain, one with more effective connections, flexibility and speed, according to our third postulate, is likely to have a shorter collapse time and thus produce more cognitive quanta in a fixed interval of time.

The usual intuition tell us that creating a more complex assembly takes more time or at that least allocating more time for construction should not undermine the soundness of such an assembly. But, the usual intuition might be wrong with the particular space-time assembly which is our subjective experience. The cognitive reasoning behind the third postulate is that a "more powerful" cognitive system should be able to perceive the existence of a "weaker" one no slower than a weaker one becomes aware of it. We cannot will our consciousness to slow down in the

process of its own creation (since to have the will, we must first achieve consciousness) nor can any external factor slow it down without first reducing its power (for example by reducing the number of active neurons and thus first *weakening* it). A more powerful cognitive system, if it remains more powerful (that is evolving at higher information density) using all its resources, must and will achieve its collapse time faster in general.

Achieving consciousness is a *phase transition*. The second postulate states that the region of complexity in terms of information density and volume, the region where we make a transition to the phenomenon of consciousness requires certain minimal concentration of information patterns in space-time. The minimal density requirement prevents the brain from being turned off and on; consciousness cannot be disrupted.

Consider the brains of two identical fish, created simultaneously at the same initial state and working concurrently. They will achieve consciousness at about the same time. Now suppose we could combine the resources of the two fish brains with (minimal) modifications. The third postulate implies that if we would ever succeed to combine the resources effectively, the total network must have a smaller collapse time: Time (fish1*fish2) < Time (fish1) =Time (fish2).

More generally our *law of cognitive composition* for two systems S and W relates the collapse times by [1]

$$T\ (S*W)\ \leq\ \min\ \big(T(S),\ T\ (W)\big).$$

How can a system with more components spread over a larger region of space keep up with additional burden of coordinating more processes and achieve consciousness even faster? The answer is faster communication. Thus, the third postulate predicts that a more powerful cognitive system must have faster communication channels. Because faster communication would probably allow the same computational component to create more complex patterns even without additional resources, the third postulate implies speed is the key in the evolution of cognitive systems.

The third postulate says a more powerful conscious experience cannot be achieved by utilizing time as a resource. This limits how powerful a cognitive system can get (by demanding temporal collapse to be shortened). In combining the brains of the two fish, thinking of them as irreducible

[1] Remarkably, this relation, arrived at by purely cognitive arguments, holds for chromatic number (instead of the collapse time function T) of tensor products (instead of cognitive composition '*') of two graphs (instead of two cognitive systems). Chromatic number of a graph is related to optimal scheduling problems [8], optimal register allocation for compilers and pattern matching. If we view an edge as a "causal link" between two "event nodes" in the information pattern of the brain, then two neighboring nodes cannot have the same color, i.e.; happen at the same time. The tensor product copies the first graph into each node of the second, and so it suggests that the sequence (of compositions) gives the pattern a fractal structure. We wrote that the collapse time is somehow minimal; the chromatic number, interpreted in this way, is the minimum time to label such a graph.

fundamental units, we cannot slow down the speeds of the two brains, we have to get even faster communication channels to coordinate and combine the two.

But, one cannot unify cognitive components whose *event influence graphs* do not have internal unified properties already. It is not believable that brain activity could exhibit unified properties at the highest space-time scales of cerebral cortex while unified properties are absent when we look at smaller sections and sub-sections of the brain. This requirement for unification at smaller scales strongly suggests that the information patterns of the brain are self-similar.

We are able to maintain consciousness under the influence of even powerful drugs and moderate mechanical force. The state of consciousness might change but unification does not break down. This means, within some limits, information patterns of the brain have topological properties and are invariant under certain transformations.

We have suggested that a cognitive system is highly connected but we cannot connect everything, every component and process to every other component and process at the same level of intensity. The key in solving the binding problem is the discovery of laws (geometric in nature) which combine space-time proximity of cognitive activity with aggregation and transfer of information. A group of neurons firing in unison in a small region of space producing an aggregate output (aggregate in the sense of information irreversibility meaning we cannot decompose the aggregate signal uniquely back to its components) can be unified with another group if this aggregate signal influences their activity within the collapse time. The influence signals could be complex wave-packets which retain some information about the individual components or weak simple signals.

The binding problem should depend crucially to *space-time-influence distance* of events. The abstract *event influence graph* is created by information signals from parts to parts, from whole to parts and from parts to whole (Information from a single neuron could be replicated and act upon whole regions of the brain). Dimensional proximity in the *space-time-influence* metric cannot be too weak in any of the three components. We cannot bind events close in spatial dimension which are days apart, no matter how strong the path of influence between them might be. The three postulates do not appear to shed much light on the geometry and properties of this metric. This work remains incomplete in other aspects as well. For example, the temporal collapse might be subject to uncertainty laws and involve some properties of the smoothly varying window over space-time dynamics mentioned earlier, i.e.; the collapse time might not have a "sharp edge".

References

[1] R. Penrose, The Emperor's New Mind. Oxford, Oxford Univ. Press, 1989.

[2] R. Penrose, in The Large, the Small and the Human Mind, ed. M. Longair, Cambridge, Cambridge Univ. Press, 1997.

[3] Max Tegmark. Importance of quantum decoherence in brain processes, Phys. Rev. E 61, 4194–4206, 2000.

[4] Tim Palmer. The Invariant Set Postulate: a new geometric framework for the foundations of quantum theory and the role played by gravity. Proceedings of the Royal Society A 465, 3165 3185, 2009.

[5] T. Metzinger, T. (ed. 2000a), Neural Correlates of Consciousness: Empirical and Conceptual Questions, Cambridge, MA: The MIT Press/A Bradford Book.

[6] Lewis H. Ryder. Quantum Field Theory, Cambridge University Press, Second Edition. 1996.

[7] Terence E. Stuart, Joshua A. Slater, Roger Colbeck, Renato Renner, and Wolfgang, Experimental Bound on the Maximum Predictive Power of Physical Theories. Phys. Rev. Lett, 109, 020402; July 9, 2012.

[8] Marx, Dániel, Graph coloring problems and their applications in scheduling, Periodica Polytechnica Ser. El. Eng. 48(1-2):5-10, 2004.

[9] Judy Kupferman, Benni Reznik, Entanglement and the Speed of Evolution in Mixed States, Phys. Rev. A 78, 042306 (2008)

[10] R. VanRullen, C. Koch, Is perception discrete or continuous? Trends Cogn Sci 7:207–213 (2003).