

# Thoughts on Qualia for Machines

Nordin Zakaria  
nordinzakaria@gmail.com

## Abstract

I speculate upon the idea that qualia comprise of quanta or packets, and that each packet is generated by physical processes within a neuron, possibly at the quantum level. Pattern-specific neuronal activation causes wave-like interactions among the packets, leading to phenomenal sensation. In essence, I provide in this paper a new panpsychic interpretation of the hard problem of consciousness.

## 1 Introduction

The past decade has seen an emerging of machine consciousness as a multidisciplinary, scientific discipline. The new discipline is driven explicitly by the desire to make the leap from the current state of art - robots and machines with “standardized and well-defined” (Sanz, 2005) interaction capability - to one with true human-like behavior. The implicit driver is philosophical - “the desire to make a machine in ones own image” (Lansky, 1996) in order to understand oneself.

We are still major leaps away though from machine consciousness. In spite of fundamental progress in neuroscience, there is much more ground to cover. Neuroscience goes in into the actual brain, studying brain circuitry and correlating activations to cognitive tasks. The work involved is akin to examining a computer system, with zero prior knowledge about computer hardware and software, to reverse-engineer its inner working. One would observe the wiring, the details of the parts available, the electrical signals made when one clicks a user interface widget, and try to infer how it all works. It is a difficult task, and the primate brain is of course much more complex than a computer. Then again, it is not an impossible task - “easy problem” as per (D. J. Chalmers, 1995). Though it will take some time, given the 100 billion neurons in the brain, it can be done with scientific will (and grants). Meanwhile, the “hard problem” (D. J. Chalmers, 1995) of consciousness, where the question of the raw senses or qualia is concerned, to this day remains largely unresolved. These are the inner variables of a conscious being, and as also noted by (Fingelkurts, Fingelkurts, & Neves, 2012), it is unlikely we will ever be close to a conscious robot until we crack nature’s code for it.

In this sort paper, I speculate upon an idea that I feel is of relevance to efforts in solving the hard problem. My contention is that qualia comes in quanta or

packets, and that each packet is generated by physical processes within a neuron, possibly at the quantum level. Pattern-specific neuronal activation causes wave-like interactions among the packets, leading to phenomenal sensation. I elaborate further in the next section.

## 2 About Qualia

What exactly are these “little wisps of protoplasm in the brain” (Ramachandran & Hirstein, 1997) known as qualia? (D. J. Chalmers, 1995) refer as qualia “those properties that characterize conscious states according to what it is like to have them”. In simpler terms, qualia refer to the raw senses; the red when I see an apple, the pain when a needle prick into my finger, and the luxurious taste of warm chocolate upon my tongue.

Qualia serve as internal variables. It dictates the state of being, and this state determines our motivation, desire and action - the captivation by an inspirational painting, the quick twitch and withdrawal of the finger upon reaching a pain threshold after being pricked, and the salivating and the desire to consume more of a particular indulgence. This cannot be just a simple variable, as it is likely to be in today’s computer implementation of cognitive architecture. Typically, in such programs, there would be variables declared that represent some “qualia” value, for example, the value of the color perceived, as illustrated below:

```
int color [10];
color[3] = 0.9;
if (color[3] > 0.5)
    curr_color = "blue";
```

Admittedly, the code above is grossly oversimplified. But the point is to show that the piece of code feels no color, even if the computer running it displays functional behavior corresponding to the recognition of color. It is not conscious, and the functional appearance of senses such as pleasure and pain has to be explicitly programmed.

The difficulty has, as said earlier, been in elucidating qualia down to its physical or structural form. We may know well the ins and outs of the physiology of the basic senses such as vision but we still do not know where in the pathways do the qualia arise. Hence a typical neuroscience book would go in depth into the anatomy and pathway involved in color vision, and would probably deal with trichromatic or opponent-process theory of color, but after all the complication and details, it would still say nothing on the question of how color itself arise.

Other than explaining it away using emergentism or functionalism, the only way to explain qualia seems to be panpsychism. As put by (D. Chalmers, 2011), we can understand panpsychism as the “thesis that some fundamental physical entities have mental states”. He said “if quarks or photons have mental states, that suffices for panpsychism to be true, even if rocks and numbers do not have mental states”. The problem with this conception is the vagueness; what is mean by mental states? Is it really all in the quarks the feel of what it is like -

the taste of meat, of spices, the sound of Beethoven, the pain of needle, and so on?

Let us now attack the question by focusing on just color. What actually is color? In the physical world, colors are but wavelength-specific reflections from the surfaces of objects. At the quantum level, light may be considered as a stream of photons that when in collision with matter impart energy to electrons contained within. The resulting change of energy in the electrons caused the absorption of light of some wavelength and the reflection of light of other wavelength. But there is no color in the physical description of the process.

Should colors be seen instead as codes in a certain language that tell us the wavelength or energy of photons that reaches our eyes? In other words, it is physical information being transmitted to our brain and decoded as colors, analogous to radio wave getting converted to electrical signals and then to sound wave and light in televisions. Still, it does not answer the what and how of color. It is indeed strange - colors do not exist in the real world. It is not out there. It only exist in human brains and that of certain animals. The question is then again: If colors do not exist in the real, physical world, and yet the brain itself is a physical device, then how can we possibly see color? My thesis is that it cannot be mere "perception"; it cannot merely be some trick of the brain.

It draws upon me that color ought to be an "inner" physical property. What meant is that it exist within physical processes, possibly down at the quantum level, within individual neurons. Color might be the inner property of electron energy shift. The reason we do not see it as existing in the external world is simply that it is an *internal* process. All of physics as we know it is concerned with observation of natural processes from the outside. We make atoms collide and measure its consequences. We make a ball drop and observe its bounce. We shine light at metallic surfaces and measure the resulting emission. All of what observed has been observed from the outside of the processes. But the brain and all of its neurons are at the very fundamental level made up of atoms, electrons and quarks too. Hence, physical processes happen within it and these processes are part of us. The inner physical processes must be occurring in neurons and other brain cells.

The same must be the case for sound, touch, taste and other qualia - these should all be inner properties. But qualia such as taste and emotion are complex. It is awkward to think of electrons or quarks having "mental" state and being able to decode the taste of chocolate. It is more useful to think of these properties as being each a composite comprising of simpler unit mental states, an idea that concur with that in (Loorits, 2014; Koko, 2013). Hence, qualia is not to be considered as a continuous irreducible quantity. I contend that each comes in "packets", analogous to light. Qualia packets interact, through concurrent neuronal activations, resulting in more complex qualia or in a complete mental state. In fact, it is conceivable that different complete qualia - eg. color, shape, etc. - are combined or "binded" to form a single perception through neural synchrony. Hence, as also noted in (Ramachandran & Hirstein, 1997; Crick & Koch, 1998; Loorits, 2014; O'Brien & Opie, 1999), the self arises from

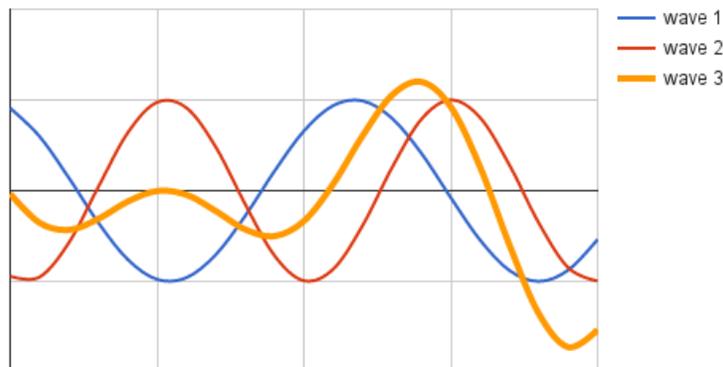


Figure 1: Qualia generation

neuronal activities.

Consider now the question of how qualia packets would interact with one another. To address this question, it is useful to think of each qualia packet as a wave-like structure. Interaction between packets then is akin to wave superposition, as shown in Figure 1 where *wave 1* and *wave 2* combine to form a more complex *wave 3*. The combination problem (D. J. Chalmers, 2012), the question of "how do the experiences of fundamental physical entities such as quarks and photons combine to yield the familiar sort of human conscious experience that we know and love" therefore should not arise. An arbitrarily complex wave can be formed from the superposition of unit waves.

So what could be the nature of the physical processes that contain qualia packet? The most successful idea is arguably that of (Hameroff, 2007), that these physical processes are quantum events in microtubules within neurons. (Hameroff, 2007) argued that quantum coherence happens in the brain microtubules, and that this coherence holds until the differences in mass-energy distribution among superpositioned tubulin states reaches a threshold related to quantum gravity. Subsequently, in a process known as Orchestrated Objective Reduction (Orch OR), the wave function describing the coherence self-collapses, creating a moment of consciousness. According to (Hameroff, 2007), a sequence of Orch OR events constitutes a "stream of consciousness".

Aside from Orch Or, there can be other quantum-level explanations, one that does not involve wave collapse, or the physical processes can even be electromagnetic in nature (Fingelkurts et al., 2012; Mcfadden, 2002). Then again, I shall not delve deeper into this aspect in this paper. My main contention, as said earlier, is that qualia is the inner properties of some sort of physical processes within the brain cells.

### 3 Discussion

In this section, I first examine the impact of the proposed idea on the axioms of natural dualism (D. J. Chalmers, 1995) as stated by (Blamauer, 2002). The axioms were as follows:

1. Physical properties are taken as fundamental properties of reality.
2. Consciousness is a fundamental and irreducible fact of reality.
3. Our universe consists only of one kind of stuff. This stuff has two fundamental kinds of properties: physical properties and mental properties.
4. These properties correspond to a set of fundamental laws that correlate the two kinds of basic properties with each other.
5. The fundamental features of reality must be ubiquitous.

My contention is with the second and third point in the axiom. Though consciousness is fundamental, it is compositional and structural, comprising of units that are in essence physical properties. The second and third point in the axioms should therefore be rewritten as follows:

2. Consciousness comprises of discrete packets, each generated by specific neurons.
3. Our universe consists only of one kind of stuff. This stuff has two fundamental kinds of properties: external, observable properties and "inner" properties.

Next, it should be noted that my conception of qualia is compatible with the three laws of qualia proposed in (Ramachandran & Hirstein, 1997):

1. *Qualia are irrevocable* - a neuronal activation pattern generates the same qualia whenever it is invoked, an assertion which is in line with the principle of organizational invariance (D. J. Chalmers, 1995). Further, this law implies that the notion of qualia being "private, subjective, and unsharable properties belonging exclusively to a private self" does not hold (Ramachandran & Hirstein, 1997).
2. *Qualia do not always produce the same behaviour* - behavior selection is the task of higher-order consciousness that takes in qualia as its input;
3. *Qualia endure in short-term memory* - this results from the causality relation between neural activation and qualia.

Finally, I consider the implication of the proposed idea on the burgeoning new endeavor of whole brain emulation. In whole brain emulation (see (Sandberg, 2013)), the basic idea is to take a particular brain, scan its structure in detail, and construct a software model that when run on appropriate hardware will behave in essentially the same way as the original brain. Since,

as I have asserted earlier, qualia are produced by specific neuronal activation via "inner" physical processes, the hardware to be used in whole brain emulation should then comprise of units and structures similar to that in the brain. Hence, assuming qualia as a quantum-based inner property, some form of quantum computer, perhaps even a connectionist quantum computer, will be needed.

## 4 Conclusions

In this paper, I have speculated on a possible explanation for the phenomena of qualia. I contend that qualia are the inner properties of physical processes within neurons, possibly at the quantum level.

Hence, my position is that panpsychism is real, but instead of associating a mental state with individual physical process, I associate with it a qualia packet or a mental unit. A mental unit is simply an inner or hidden physical property that we cannot observe from the outside. We can only "experience" it when units combine to form a mental state. But why should this be even true? As noted in (Velmans, 2008), this would be the real "hard" problem, in the same sense if we were to ask, for example, why should electricity be coupled with magnetism and why should particle-wave duality be true.

## References

- Blamauer, M. (2002). Taking the hard problem of consciousness seriously : Dualism , panpsychism and the origin of the combination problem. In M. Blamauer (Ed.), *The mental as fundamental;new perspectives on panpsychism*.
- Chalmers, D. (2011). Panpsychism and panprotopsychism. *The Amherst Lecture in Philosophy*, 1–32.
- Chalmers, D. J. (1995). The Puzzle of Conscious Experience. *Scientific American*(December), 90–99.
- Chalmers, D. J. (2012). The Combination Problem for Panpsychism. (July 2012), 1–32.
- Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, 8(2), 97–107.
- Fingelkurts, A. a., Fingelkurts, A. a., & Neves, C. F. H. (2012). "Machine" consciousness and "artificial" thought: An operational architectonics model guided approach. *Brain Research*, 1428, 80–92.
- Hameroff, S. (2007). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *NeuroQuantology*, 5(1), 1–8.
- Koko, J. (2013). *Arguments and Model for Quantum Consciousness , Modification of Quantum Collapse , and Panpsychism* (Tech. Rep.). Retrieved from [vixra.org/abs/1306.0142](http://vixra.org/abs/1306.0142)
- Lansky, A. L. (1996, June). *Consciousness as an active force*. Retrieved from <http://www.renresearch.com/consciousness.html>

- Loorits, K. (2014). Structural qualia: A solution to the hard problem of consciousness. *Frontiers in Psychology*, 5(MAR), 1–9.
- Mcfadden, J. (2002). The Conscious Electromagnetic Information (Cemi) Field Theory The Hard Problem Made Easy? *Journal of Consciousness Studies*, 9(8), 45–60.
- O'Brien, G., & Opie, J. (1999). A connectionist theory of phenomenal experience. *The Behavioral and brain sciences*, 22(1), 127–148; discussion 148–196.
- Ramachandran, V. S., & Hirstein, W. (1997). Three Laws of Qualia. *Journal of Consciousness Studies*, 4(5), 429–458.
- Sandberg, A. (2013). Feasibility of Whole Brain Emulation. *Philosophy and Theory of Artificial Intelligence*, 5, 251–264.
- Sanz, R. (2005). Design and implementation of an artificial conscious machine. In *Iwac2005*.
- Velmans, M. (2008). How to separate conceptual issues from empirical ones in the study of consciousness. In *Models of brain and mind* (pp. 1–11).