Chalmers' principle of organizational invariance makes consciousness fundamental but meaningless spectator of its own drama

Danko Dimchev Georgiev, MD

Division of Electron Microscopy, Medical University, Varna 9000, Bulgaria
Department of Emergency Medicine, 3, Bregalnitsa Street, Varna 9000, Bulgaria
E-mail: dankomed@softhome.net

Abstract

It is argued that if consciousness is a fundamental ingredient of reality then no any psychophysical law such as Chalmers' principle of organizational invariance is needed to keep coherence between experience and function (conscious action). Indeed Chalmers' proposal suggests epiphenomenal consciousness and is regress to a nineteenth century absurd philosophy. The quantum mechanics is the most successful current physical theory and can naturally accommodate consciousness without violation of physical laws.

The hard problem

David Chalmers, a well-known philosopher working on theory of consciousness from University of Arizona, Tucson, AZ, has developed the arguments proposed first by Nagel (1) that experience is *the hard problem* that makes *the mind-body problem* intractable. Functionalist's theories of consciousness, in general, are not able to explain why performing of given function must be associated with any experience. In the materialistic school the current paradigm claims that the *sentient brain* emerges at critical level of complexity from *insentient matter*. However because the chemical atoms are essentially identical both for the *sentient brain* and the *insentient matter* we can start 'recipe' for building up a sentient brain from insentient matter and we must have answer when our product will '*start to feel*'. If we follow such recipe for building up a sentient brain we will come to a situation in which adding or removing a single atom will '*give*' or '*take*' the mind of the system - conclusion quite intolerable for the ordinary logic. Thus we came to the conclusion that experience must be fundamental ingredient of the Universe we live in, that there is no insentient matter at all, and that we must *'construct'* sentient brain from sentient matter. What is interesting for this unordinary conclusion is that it is quite frustrating for us but in the same time it provides us with insight how to better understand ourselves!

The principle of organizational invariance

Chalmers (2,3) derives the same conclusion that consciousness (experience) must be a *fundamental ingredient of reality* but further he tries to formulate *psychophysical laws* that are linking experience with physical processes. Thus the principle of organizational invariance is born - a rather unsuccessful solution that hides in itself the seeds of its own destruction. In order to understand the failure of Chalmers' classical approach let us examine the following neural replacement scenario (4-7); extended citation from Chalmers (3):

"We can imagine, for instance, replacing a certain number of my neurons by silicon chips. In the first such case, only a single neuron is replaced. Its replacement is a silicon chip that performs precisely the same local function as the neuron. We can imagine that it is equipped with tiny transducers that take in electrical signals and chemical ions and transforms these into a digital signal upon which the chip computes, with the result converted into the appropriate electrical and chemical outputs. As long as the chip has the right input/output function, the replacement will make no difference to the functional organization of the system. In the second case, we replace two neighboring neurons with silicon chips. This is just as in the previous case, but once both neurons are replaced we can eliminate the intermediary, dispensing with the awkward transducers and effectors that mediate the connection between the chips and replacing it with a standard digital connection. Later cases proceed in a similar fashion, with larger and larger groups of neighboring neurons replaced by silicon chips. Within these groups, biochemical mechanisms have been dispensed with entirely, except at the periphery. In the final case, a chip has replaced every neuron in the system, and there are no biochemical mechanisms playing an essential role. We can imagine that throughout, the internal system is connected to a body, is sensitive to bodily inputs, and produces motor movements in an appropriate way, via transducers and effectors. Each system in the sequence will be functionally isomorphic to me at a fine enough grain to share my behavioral dispositions. But while the system at one end of the spectrum is me, the system at the other end is essentially a copy of silicon robot".

Now come the questions:

> *Will this silicon robot be conscious?*
>
> *If not, what about the intermediate cases: do qualia suddenly disappear or they are slowly fading?*

Further analyzing the behavior of the system that is half-robot/half-human and arguing that qualia neither are fading, nor suddenly disappear Chalmers (3) erroneously concludes that *any functional isomorph of a conscious system must have qualitatively identical experiences and the silicon robot must be conscious too*. But what indeed means the principle of organizational invariance? It can easily be proved that this principle makes consciousness epiphenomenon, therefore causally ineffective!

Proof: *Since we speak of functional isomorphs then the functional organization and the outputs of the systems must be primarily defined and the both systems must have the same functional organization and outputs. Then Chalmers' principle ascertains us that both systems must have identical experience! But in a fully determined functional system any additional ingredient must not have any causal power in order the system to preserve its functional identity. The added experience to both isomorphs then will be fundamental but causally ineffective i.e. epiphenomenon. Q.E.D.*

Following the above conclusion we can be pretty sure that Chalmers' principle of organizational invariance makes consciousness fundamental but meaningless spectator of its own tragedy. We face up the questions:

> *What is the use of such tragic/parodic universal entity?*
>
> *If experience is a fundamental ingredient of reality do the current physical theory offer solution to the mind-body problem?*
>
> *Can consciousness escape the epiphenomenal philosophy?*

Epiphenomenalism is an absurd philosophy

The epiphenomenalism is a sort of one-way dualism, in which consciousness is a product of brain processes but is itself without any causal effect on those processes. The central motivation for epiphenomenalism lies in the premise that all physical events have sufficient causes that lie within the class of physical events. If a mental event were something other than a physical event, then for it to make any causal contribution of its own in the physical world would require a violation of physical law. Huxley (8) who held the view compared mental events to a steam whistle that contributes nothing to the work of a locomotive. Taking into consideration that the historic roots of epiphenomenalism trace back to nineteenth century it is important to

note that under *physical* in the above statement is meant *material* (in the following discussion we will revisit both terms).

Nevertheless the epiphenomenalism is absurd; it is just plain obvious that our pains, our thoughts, and our feelings make a difference to our (evidently physical) behavior; it is impossible to believe that all our behavior could be just as it is even if there were no pains, thoughts, or feelings (9). Also if our souls are attached to the bodies and just 'experience' what the body is doing without having any possibility to react, who decides which soul to which body will be attached? A Universal Marquise de Sad who decides millions of children in Africa to 'suffer' just without any sense?

The clinical practice has not shown mismatch between the experience (e) and the neural events (n) to exist. Neither *inverted qualia* (e.g. pleasant feeling associated with neural events that normally evoke pain like tachykinin neuromediation), nor *dancing qualia* (e.g. sudden and unpredictable flipping between pain and pleasure associated with neural events that normally evoke pain) have been observed. Mismatch between the objective reality and the experience is only possible if the peripheral reception is affected i.e. the stimulus elicits neural event that normally must not occur! Such mismatch deteriorates the survival capability of the organism and cannot be tolerated by the natural selection. Individuals with *syringomielia* have malformation of the spinal cord and lack feeling of pain when detrimental stimulus is applied to certain area of their extremities. These patients are not aware when noxious stimulus is applied, so often they obtain heavy injuries (wounds or burns) of their arms, legs, face or chest, not having chance to react adequately to the situation. The conclusion is that the unpleasant feelings have survival value.

Thus we face to another huge problem that epiphenomenalism must solve - the development of consciousness must be explainable through natural selection. But a property can be selected for only if it has an effect upon organisms' behavior. Therefore, consciousness (both *qualia* and *intentional states*) must have effects in behavior, i.e., epiphenomenalism is false. Today, this argument is generally associated with Popper and Eccles (10). However it is an old argument and clear statements of it were offered by James (11,12) and by Romanes (13). James offered an intriguing variant of the argument from natural selection. If pleasures and pains have no effects, there would seem to be no reason why we might not abhor the feelings that are caused by activities essential to life, or enjoy the feelings produced by what is detrimental. Thus, if epiphenomenalism were true, the felicitous alignment that generally holds between affective valuation of our feelings and the utility of the activities that generally produce them would require a special explanation. Yet on epiphenomenalists' assumptions, this alignment could not receive a genuine explanation. The felicitous alignment could not be selected for, because if affective valuation had no behavioral

effects, misalignment of affective valuation with utility of the causes of the evaluated feelings could not have any behavioral effects either. Epiphenomenalists would simply have to accept a brute and unscientific view of pre-established harmony of affective valuation of feelings and the utility of their causes.

According to Robinson (14) the above argument can be met by supposing that both the pleasantness of pleasant feelings and the pleasant feelings themselves depend on neural causes. However these neural causes will give rise to something so useless (pleasantness of the pleasant feeling) as is the pleasant feeling itself. In its turn the epiphenomenalists must invoke another ad hoc supposition in order to explain why should the pleasantness of the pleasant feeling be tolerated by the natural selection! In other words epiphenomenalism requires infinite number of ad hoc psychophysical laws.

Proof: *In epiphenomenalists' approach we have objective reality (r) causally evoking neural event (n1) and try to explain why experience (e1) is adequate to the objective reality e.g. we feel pain when detrimental stimulus elicits the neural event n1. For convenience we will assume that in general there is perfect match between the reality and the elicited neural events thus what remains to be explained is the perfect match between the experience (e1) and the neural event (n1). In order to explain why e1 is adequate to n1 epiphenomenalists assure us that this adequacy (e2) depends on neural cause too (the neural cause is either the same - n1 or another one - n2). What is missed is that in both cases e2 is epiphenomenon since the adequacy e2 per se cannot have effect on organism's behavior. What has happened is just a shift of the problem that now is at e2, namely why should e2 be tolerated by the natural selection. The perfect match between e2 and the neural events (n1, or n1+n2) must on its turn be explained. If we follow the epiphenomenalists' solution then we should propose that this adequacy (e3) between e2 and the objective reality depends on neural cause(s) (n1, n1+n2, n3, n1+n2+n3). Thus we come to an infinite regress in order to keep fundamentally false theory i.e. the adequacy between experience and the neural events is either unsolvable in epiphenomenalists' approach or requires infinite number ad hoc psychophysical laws. The conclusion is that the consciousness must be causally effective in order to fight for its own existence (survival)! Q.E.D.*

In a recent paper Chalmers (15) investigates different philosophical solutions of the consciousness problem reaching to the conclusion: *"As I see things, the best options for a nonreductionist are type-D dualism, type-E dualism, or type-F monism: that is, interactionism, epiphenomenalism, or panprotopsychism".* We already have discussed the senseless and *ad hoc* character of epiphenomenalists' approach, so Chalmers' investigation is nothing but sciolism, when it comes to epiphenomenalism. Indeed he

does not realize the parodic/tragic entity in which consciousness is transformed if one decides to invent psychophysical laws of any sort. Chalmers (15) continues:

"*Another objection holds that if consciousness is epiphenomenal, it could not have evolved by natural selection. The type-E dualist has a straightforward reply, however. On the type-E view, there are fundamental psychophysical laws associating physical and phenomenal properties. If evolution selects appropriate physical properties (perhaps involving physical or informational configurations in the brain), then the psychophysical laws will ensure that phenomenal properties are instantiated, too. If the laws have the right form, one can even expect that as more complex physical systems are selected, more complex states of consciousness will evolve. In this way, physical evolution will carry the evolution of consciousness along with it as a sort of byproduct".*

What is overlooked by Chalmers is that there must be infinite *ad hoc* psychophysical laws in order to keep the theory, because as it was shown there is infinite regress in epiphenomenalists' explanation. Even the idealists' model of ideas kept coherent by God (16) seems to require only one metaphysical God than infinite ad hoc psychophysical laws. In the science however neither the appearance of infinite regress (vicious circles) nor invoking God is acceptable!

Quantum principles instead of psychophysical laws

The current physics is on its way to solve one of the biggest mysteries: the nature of consciousness. If the argument following from the hard problem analysis is correct then consciousness is a fundamental ingredient of reality. However we have seen that introducing psychophysical laws hides the risk to turn such fundamental ingredient into epiphenomenon. Then, how we could insert consciousness at the fundamental physical level without violating the physical laws? The possibility to merge neuroscience and physics is offered from the *quantum theory*, which is 'battle tested' and for now there is no experimental evidence that it fails in its predictions. What will be final form of the quantum mind philosophy is still unknown but physicists agree that there is no need the mathematics (basic laws) of the quantum theory to be reformulated in order to accommodate consciousness, nor explaining consciousness using the quantum physics will lead to breaking of the physical laws. Indeed the quantum coherent systems manifest *free will* in their evolving in time. Such systems could be in several physical states at the same time satisfying the *principle of superposition*, and what is most important all the subcomponents are *entangled* so that the physical correlations are enforced faster-than-light. Entanglement is experimentally verified phenomenon that explains why the quantum coherent

systems behave as unitary objects i.e. the quantum system is not mere sum of its parts. The *quantum wave function* (or simply the *quantum wave* denoted as ψ) of a quantum coherent system evolving in time is *nonmaterial* physical entity and itself may represent the conscious state (experience) of the system. Here *nonmaterial* means outside the 4D space-time continuum originally proposed by Einstein while formulating the theory of relativity. That is why with the use of the quantum theory the epiphenomenalists' reasoning could be bypassed i.e. mind may be *nonmaterial* (outside the space-time) but is still *physica*l (therefore fully causal). What is more important is that the *quantum wave* ψ analogous to experience is *intrinsic* and *unobservable*, because any attempt to measure it directly collapses the wave function of the system and destroys the superposition of states. The resultant philosophy could be called *quantum dualism* since mind is *nonmaterial* (the term *physical monism* used by Chalmers is not appropriate since we can adopt by definition that everything not physical is non-existing).

Now let's turn back our attention to the neural replacement scenario by silicon chips. What we can say of the composite system that contains both neurons and silicon chips? It immediately follows that because of the different atoms from which are constructed the original neural network, the intermediate case(s) and the final silicon chip robot they all will have different quantum wave functions. Also it is not obvious that the composite system will have 'single' experience just because of the fact the quantum coherence must extend throughout the whole system and surely this will not be the case since the quantum coherence is too fragile and there must be evolutionary evolved mechanisms to protect it (17). From the introduced quantum dualistic point of view it follows that all those different systems will have different experiences contrary to Chalmers' suggestion even if all of them produce the same outputs (which of course could be impossible to achieve, but this is not relevant to our stream of reasoning).

Thus quantum principles could be enough for building up a consistent theory of consciousness and suggest that no additional psychophysical laws are needed to keep coherence between experience and conscious action. The quantum theory also suggests that consciousness (experience) is a fundamental ingredient of the physical Universe we live in (a novel paradigm that surely will change also the way we understand the moral principles and ethics!), what remains to be experimentally verified is that our minds are quantum coherent systems.

References

1. Nagel, T. What is it like to be a bat? *Philosophical Review 1974;* 4: 435-450.
2. Chalmers, D.J. Facing up to the problem of consciousness. *Journal of Consciousness Studies 1995;* 2: 200-219.
3. Chalmers, D.J. Absent qualia, fading qualia, dancing qualia. In: *Conscious Experience*, Thomas Metzinger (ed.), Imprint Academic, 1995.
4. Pylyshyn, Z. The causal power of machines. *Behavioral and Brain Sciences 1980;* 3: 442-444.
5. Savitt, S. Searle's demon and the brain simulator reply. *Behavioral and Brain Sciences 1982;* 5: 342-343.
6. Cuda, T. Against neural chauvinism. *Philosophical Studies 1985*; 48: 111-127.
7. Searle, J.R. The rediscovery of the mind. Cambridge, MA: MIT Press, 1992.
8. Huxley, T.H. On the hypothesis that animals are automata, and its history. *The Fortnightly Review 1874;* 16: 555-580. Reprinted in Method and Results: Essays by Thomas H. Huxley (New York: D. Appleton and Company, 1898).
9. Taylor, R. Metaphysics. Englewood Cliffs, NJ: Prentice Hall, 1963.
10. Popper, K. & Eccles, J.C. The self and its brain. New York: Springer-Verlag, 1977.
11. James, W. Are we automata? *Mind 1879*; 4: 1-22.
12. James, W. The principles of psychology. Holt, New York, 1890.
13. Romanes, G.J. Mind and motion, and monism. London: Longmans, Green, and Co, 1896.
14. Robinson, W.S. Epiphenomenalism. In: Stanford Encyclopedia of Philosophy, 1999.
15. Chalmers, D.J. Consciousness and its place in nature. Published in (S. Stich and F. Warfield, eds) Blackwell Guide to the Philosophy of Mind (Blackwell, 2003), and in (D. Chalmers, ed) Philosophy of Mind: Classical and Contemporary Readings (Oxford, 2002).
16. Churchland, P.S. & Churchland, P.M. Neural worlds and real worlds. *Nature Reviews Neuroscience 2002*; 3: 903-907.
17. Jibu, M. & Yasue, K. What is Mind? - Quantum Field Theory of Evanescent Photons in Brain as Quantum Theory of Consciousness. *Informatica 1997*; 21: 471-490.