THE PROBLEM OF PHENOMENAL CONSCIOUSNESS AND CONNECTIONIST REPRESENTATIONALISM

Reena Cheruvalath*

Consciousness is apparently an intractable problem for the ambitious program of a science of the mind that is known as "cognitive science". There seems to be something elusive in its nature that continually frustrates various attempts of a scientific nature to provide a tenable theory of consciousness. In this paper my aim is to highlight this elusive feature of consciousness; namely, the phenomenal feature, and to discuss recent connectionist and representationalist development that advances ways of accounting for this feature. Finally, I shall take up in particular the connectionists' "vehicle" theory for a close scrutiny, and shall argue against the tenability of this theory.

1. Two Aspects of Consciousness

There are two ways one can characterize consciousness. One is "consciousness is as consciousness does", and the other, "consciousness is as consciousness seems". The former is a *causal* characterization, in which the "causal role" that consciousness plays in the general economy of our mental lives is considered as basic. Here, the consciousness leads to certain activity or behaviour. For example, consciousness of the examination may make a student engage in his or her studies. In contrast, the latter is a *phenomenal* characterization, in which the way our mental lives "seem" to us is fundamental. That is to say, conscious states are states of something appearing somehow to a subject. It seems to one someway to see red, and some other way to hear a crash, to suffer pain or to visualize a triangle. Again, it also seems to one someway to see yellow or violet, to another way to hear music or song, and to another way to hear the cry of a baby. Phenomenal characterization is given fundamentally in first person terms, describing episodes of inner life in terms of how they feel or seem to the subject who experiences them. Some example are perceptual experiences, such as seeing, tasting, bodily sensational experiences, such as those of pains, tickles, feelings of elation, itches, imaginations, experiences

* Faculty Humanities and Management, BITS Pilani Goa Campus, Zuarinagar, 403 726 Goa, India, e-mail: reennaac@yahoo.com.

such as those of one's own action or perceptions, and streams of thought, as in the experience of thinking "in words or in images".

Similarly, consciousness can be classified as "access consciousness" (A-consciousness) and "phenomenal consciousness" (P-consciousness) (Ned Block 2002). Mental states such as occurring thoughts or judgments are access conscious, which are not phenomenally conscious. Access conscious states are poised for direct control of thought and action. P-conscious, on the other hand, is explained in terms of phenomenal qualities that we typically associate with perceptual experiences, pains, pleasure, tickles etc. Thus, phenomenal consciousness is the sort of consciousness that is integral to experiences and sensations. It is found wherever there is experience and feeling, and it requires no reflective abilities. Here, the classification is given only to clarify the concept of phenomenal consciousness, as many kinds of consciousness are available.

The causal and phenomenal kinds of consciousness, however, are not mutually exclusive, in the sense that sometimes phenomenal consciousness may act causally. For example, when one feels headache, he or she goes to take aspirin. Here, the consciousness of pain leads to certain action; pain causes behaviour. These may be two stages of consciousness. There is considerable dispute over this issue, i.e., whether mental states can be conscious in the "phenomenal kind" without also being conscious in the "causal kind". Here, I am concerned with another problem, namely whether phenomenal consciousness can be explained in cognitive or connectionist representational terms.

Phenomenally, conscious experience remains always as inexplicable in the manner of other physical entities. The basic reason is that these experiences have certain phenomenal (qualitative) properties, known as "qualia", which possess aspects of what experiences are like for their bearers. This always escapes the net of cognitivists' explanation. The word "qualia" refers to the intrinsic properties of our experience, such as that of colours, sounds and sensations of pains. These are also described as "the ways things seem to us". (The term qualia is often used interchangeably with phenomenal consciousness). Qualia are regarded as persistent obstacles to functionalist as well as physicalist explanation of conscious experience in general. The difficulty lies in the implication that qualia can neither be defined in terms of their functional role or physical make up, nor can they be captured in a functionalist or physicalist explanatory framework. That is to say, it is difficult to understand how any physical property or event in our brains could be or could realize a phenomenal state, which is intrinsic, private, ineffable, and known with certainty.

Therefore, the problem of phenomenal consciousness should be treated as a hard problem, which can be distinguished from the easy problem of consciousness. It is hard to answer the question "how do physical processes give rise to subjective experience?" because experiences are paradigmatically phe-

nomenal, characterized by what it is like to have them. On the contrary, the easy problem concerns the explanation of various cognitive functions such as discriminatory abilities, reportability of mental states, focus of attention, and control of behaviour. All these functions are straightforwardly susceptible to explanation in terms of computational or neural mechanisms. A human subject can discriminate the sensation of smell from the sensation of touch. After collecting information through various sense organs, it is integrated in the midbrain, with the knowledge that our memories store. We can also report our mental states by using language.

However, it is hard to explain phenomenal consciousness, because whenever we try to explain these in neuronal or representational terms, such explanations provide mechanisms for instantiating certain functions, which must fall short of explaining the feel possessed by many types of conscious states. Consequently, it invites an explanatory gap between mental states and physical states. It arises when we lack any bridging principle between some mental facts and the other, and also between mental facts and physical facts (Cheruvalath 2007).

## 2. Can connectionist Representationalism diffuse the problem?

Connectionism is an approach in the filed of cognitive science which explains mental states by using an artificial neural network. Neural networks are simplified models of the brain composed of large numbers of units knit together, known as "nodes", with weights that measure the strength of connections between the units. They are designed to mimic the mechanisms of the neurons in the human brain. Zachary argued that biological neurons are made up of dendrites, the soma or cell body, an axon, a terminal button, a synaptic gap and synapse or point of contact with another neuron. A neuron operates through a series of electrochemical pulses. Below is the picture of a typical neuron.
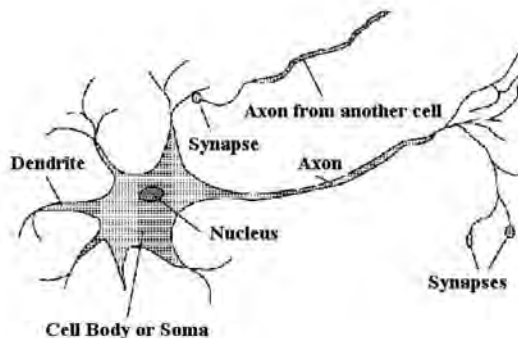
PICTURE 1

The function of a neuron is as follows: an electrical pulse is passed into the neuron through the synaptic gap, to either the dendrites or the soma. This happens until enough of the electricity amasses in order to overcome what is known as a threshold, at which the neuron fires. Firing occurs when an electrical pulse is released down the axon through the terminal button toward the synaptic gap, allowing for a message to propagate to subsequent neurons. When this structure is implemented in a computer, the basic neuron structure is simplified. In the computer representation there is no synaptic gap, and the terminal buttons are connected directly to the soma.

A model of artificial neural network is given below.
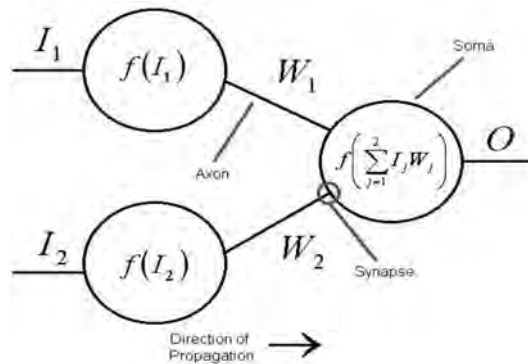
PICTURE 2.

(In this model f is some activation function and W_1 and W_2 are the connection weights).

Neural networks are represented in computers as weighted and directed graphs in which the input is entered at the extreme left through the input neurons and propagated through the network towards the right. The sum of all the input weights into a neuron, multiplied by the values of its previous neighbours, is passed to a firing function known as the activation function. The result of this function then determines the firing of the current neuron.

Recently, the connectionists have developed a new theory, known as Representational Vehicle Theory (RVT) by which phenomenal consciousness can be explained in terms of the representational vehicles or nodes that encode information in the brain (O'Brien, Gerard and Opie 1999). Their major aim is to remove the pitfalls of classical cognitivism. In the classical cognitive theories, subjective experience or phenomenal consciousness is explained either in terms of information processing system, or its representational content (Fodor 1975, Baars 1988, Michael Tye 1995). In other words, consciousness corresponds to a limited-capacity processing system and it appears as a result of

the rich and widespread informational access relations possessed by a relatively small subset of the information – bearing states of a cognitive system (Baars 1988, Johnson Laird 1988, Dennett 1991).

Different from classical cognitivism, the RVT claims that phenomenal experience is to be explained not in terms of what neural representations *do*, but in terms of what they *are*. According to the latter, phenomenal consciousness is identical with *explicit* representation of information, i.e., stable pattern of information in PDP network realized in the head. It means that a neural network realizes a stable pattern when its constituent neurons are firing simultaneously at a constant rate. It is *explicit* in the sense that each distinct item of information in a computational device is encoded by a physically discrete object.

Obviously, this reflects the representationalist principle that mental activities are entirely a matter of their representational content, i.e., a matter of what those experiences represent the world as being. On this view, for example, to say that my colour experience of red things has a distinctive qualitative character is just to say that those experiences represent such things to me as having a certain property, namely redness. So, the difference between an experience of green and an experience of red will be explained as a difference in the properties represented-reflective properties of surface, in each case. Here, it means phenomenal experience can be represented as neuronal vehicles.

To overcome the difficulties faced by representationalists, RVT suggests that the neural vehicles of explicit representations have certain intrinsic properties. In other words, network activation patterns are physical objects with intrinsic structural properties just as much as neurotransmitter molecules, neurons, and brain. Thus, for the connectionists, consciousness is an intrinsic physical intra-network property. Informational access relations, which depend on the capacity of each network to have effects on the others, inhere in these collections of interconnected network. In these relations, stability of activation has an important role. That is to say, information processing is not complete until a stable pattern of activation has been achieved. According to them, the internal structure of conscious experience is determined by relations of "structural isomorphism" between network activation patterns and certain properties of the world. Thus, a marriage of the "vehicle theory of consciousness" with the "structural isomorphism theory" will help to close the explanatory gap. Two events are structurally isomorphic in the sense that the features of the former resemble the latter. In other words, the conscious experience is structurally isomorphic to the neural network.

However, it is easy to conceive of a non-human system in which informative representations are distributed to all subsystems, yet those representations are totally devoid of phenomenal awareness. A non-human system coor-

dinates its inputs and processes available data without having the feeling that it is processing. As John Searle in his Chinese Room argument said, system does not have understanding (Searle 1999). Anyhow, RVT leads to the fact that phenomenal conscious would correspond to some system level property, i.e., information processing characteristics of neurons.

3. CHALLENGES TO RVT

More generally, connectionists' support of Representationalism invites all the hurdles faced by the representationalists. First, informativeness is not sufficient for phenomenal consciousness. It may be formed by the operational process, but the process analysis does not disclose much about the *states* formed. Secondly, RVT suggest that subjective experience can be identified with explicit representation of information. This raises the question do why the information explicitly represented in the brain is conscious when the information explicitly represented elsewhere, for example, on this sheet of paper, is not. It shows that a stable pattern of activation is not sufficient to generate phenomenal experience. The theorists themselves admit that the explicit information can occur even in the absence of conscious experience since there are unconscious explicit representations in human cognition. Besides, it also depends on other factors, like chemical or physical processes in the brain. Again, there are many phenomenal experiences which cannot be explicitly represented, for example, moods.

Moreover, if we identify the feeling of pain with this stable pattern of activation, we would have given a reductive explanation of pain, but the reduction of pain to this pattern still leaves the subjective experience of pain unreduced. That is to say, even if we accept that phenomenal consciousness is a stable pattern of activation, how conscious experience generates or emerges from this vehicle and closes the explanatory gap is still a problem (Seager 2000). Thus, the connectionist effort of closing the explanatory gap is doubtful.

The connectionist's suggestion that a marriage of the vehicle theory of consciousness with a structural isomorphic theory provides a solution does not seem to be feasible. The "internal structure" of conscious experience is determined by the relations of structural isomorphism between network activation patterns and certain properties of the world. This brings about another difficulty. One might know everything about the "shape" and "dimensionality" of a given neural activation space, yet it is impossible to know *what it is like* to have the corresponding experience. For instance, if we arrange a net to operate in a non-human system, with an "activation space shape", which is quite unlike that of human sensory modalities, then we cannot know what such an artificial network would experience. It means that if we can know the "shape" of the space and still not know what it is like to have the experience, then having a particular activation pattern cannot be all that there is to

having an experience. Therefore, connectionists lack the resources to close the explanatory gap. Although, connectionist's position is promising, its feasibility is uncertain, because, these kinds of constructs are well suited to the easy problem of consciousness alone, and not to the hard problem, because, as I have mentioned earlier, to explain a cognitive function, we need only to specify the mechanisms which can perform the function. Nonetheless, in the hard problem, the problem persists even when the performance of all the relevant function is explained.

If each feeling or experience has particular representational content, then again, representationalists have to face the problem of identifying the relevant *representata*. For example, sensation of heat represents neither heat, nor temperature nor conductivity, nor energy (Gray 2003). That is to say, the sensation of heat is different from heat or temperature. The trouble is that heat, temperature or energy may exist without being felt as heat, i.e., these may exist without producing the sensation of heat. Here, the representational content is different from the phenomenal content. Similarly, if consciousness is represented as "nodes" or neurons, the RVT theorists will be in trouble to find out which node represents which feeling. We have different degrees of feeling, i.e., the intensity varies.

Likewise, representational content may exist without having any subjective experience. For instance, I have a wound on my right hand and I feel pain. Then, in representationalist terms, pain has some representational features in a way that a pain is felt as being in a certain part of one's body, as if that part is disordered in a certain way (Tye 1995 & Lycan 1996). That is to say, the pain can be explicitly represented on my right hand. However, sometimes, it happens that when one of my close friends visits me, for a while I forget my pain or I do not feel pain at all. Accordingly, the representational content is there without any feeling. In the same way, if I have a similar kind of wound (having the same length and breadth) on my left hand also, how can I differentiate these two kinds of pain? Thus, it shows that there exists a gap between the representational content of experience and its phenomenal character.

Perhaps, connectionists should be appreciated for using different interdisciplinary methods to solve the problem of consciousness. In the same way, they can provide an account of all components of mental life and also to explain how these components help to cause certain behaviour. It is difficult however to capture its qualia, lacking the theoretical tools to settle decisively questions about whether any two functionally equivalent systems differ in their phenomenal aspects, or even whether a given system has any qualia at all. Therefore, the explanatory power of any mechanistic or representationalistic or, in general, connectionist schemes to account for the presence or emergence of qualia is doubtful. No physical mechanism seems very intuitively plausible as a seat of qualia, least of all a brain. Since we know that we are brain-headed

system, and that we have qualia, we know that brain headed systems can have qualia. Moreover, behaviour, functional role and causal relations are irrelevant to the existence of conscious mental phenomena. The reason is that while the functional or behavioural evidence may establish intentional similarities and differences between the (mental states of) subjects, it cannot capture qualitative similarities and differences.

Here it seems, by identifying phenomenal consciousness with vehicles of information processing system, connectionists try to build in terms of neuronal functions. In this sense, vehicle theorists are not very different from the classical cognitivists. In other words, it signifies only what neuronal representations *do,* not what they *are.* The goal of such an account is to characterize consciousness as "causal kind" and not as "phenomenal kind". Certainly, this shows that the vehicle theory can solve the easy problem of consciousness. At present, in the relevant sciences we have no idea how the neural substrate of my pain can explain why my pain feels like this rather than some other way or no way at all. However, though the intricacies of brain function may be opaque to us now, it does not mean that it will remain so forever. Besides, if there is any hope of diffusing the mysteries of phenomenal consciousness, it would have to depend on an interdisciplinary, multi-perspective approach hitherto not complete.

REFERENCES

1. BAARS, B., *A Cognitive Theory of Consciousness*, Cambridge University Press, Cambridge 1988.
2. CHERUVALATH, R., *Explanatory Gap and Mental Causation*, in «Teorema-International Journal of Philosophy», vol. XXVI/I (2007), pp. 107-115.
3. DENNETT, D., in ALLEN LANE, *Consciousness Explained*, The Penguin Press 1991.
4. FODOR, J., *The Language of Thought*, Harvard University Press, Cambridge MA 1975.
5. GRAY, R., *Tye's Representationalism: Feeling the Heat?*, «Philosophical Studies», 115 (2003), pp. 245-256.
6. JOHNSON-LAIRD, P. N., *A computational analysis of consciousness*, in A. MARCEL, E. BISIACH (eds.), *Consciousness in contemporary science,* Clarendon, Oxford 1988.
7. KENDALL, G., "Neural Networks", *Introduction to Artificial Intelligence*, 21 September 2001.
8. NED BLOCK, *Some Concepts of Consciousness*, in *Philosophy of Mind: Classical and Contemporary Readings*, D. CHALMERS (ed.), Oxford University Press, Oxford 2002.
9. O'BRIEN, G. and O., *A Connectionist Theory of Phenomenal Experience*, «Behavioural and Brain Sciences», 22 (1999), pp. 127-148.
10. ROJAS, R., *Neural Networks: A Systematic Introduction*, Springer-Verlag, New York NY 1996, pp. 18-19, 144.
11. SEARLE, J., *The Chinese Room*, in WILSON R.A, and F KEIL (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, MIT Press, Cambridge MA 1999.

12. Tye, M., *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind,* mit Press, Cambridge ma 1995.

13. Lycan, W. G., *Consciousness and Experience*, Bradford Books, Cambridge ma 1996.

14. Zachary, A. K., *Associative Memory and the Board Game Quarto Cross Roads*, «The acm Students Magazine», Issue 10. 2 (Winter 2003), Cognitive Science.

abstract: *Subjective experience, or the phenomenal aspect of consciousness, apparently is an intractable problem for the ambitious programme of the science of the mind that is known as "cognitive science". Connectionism, one of the approaches in the field of Cognitive science, tries to explain consciousness in cognitive or representational terms. This sort of representationalism faces a formidable obstacle. It leaves out the subjective experiential quality of consciousness, in that any reductive explanation of the subjective quality of experience in neurological, cognitive or functional terms is destined to miss out on the fact of the feel of a conscious state. Neural constitution and function are a kind of fact different from, even though related to, the kind of fact that a subjective experiential feel occurring in a neuro-functional condition. This in a nutshell, is the problem of phenomenal consciousness. However, Connectionist cognitive scientists profess the view that a solution to this problem is possible through what they describe as the Representational Vehicle Theory (rvt). rvt identifies phenomenal consciousness with stable pattern of activation or explicit neural representation of information. In this paper, I shall argue that rvt might only be adequate to solve the "easy" problem of consciousness, and not the "hard" problem. For it is designed to specify the mechanism of cognitive functions underlying subjective experience, whereas the problem of subjective feel persists even after the performance of all the relevant functions are explained.*