Proceedings of the 4th International Conference on Cognitive Science (Sydney, July 2003) pp.439-443

The easy and hard problems in cognitive science.

Ken Mogi (kenmogi@csl.sony.co.jp**)**
Sony Computer Science Labratories
Tokyo, Japan
http://www.qualia-manifesto.com

Abstract

Chalmers (1995) isolated the so-called "hard problem" of consciousness from the "easy problems". His thesis did the research community a great service by pointing out correctly that the problem of consciousness cannot be just swept away with a functionalist or connectionist description. Here, citing some recent studies from cognitive neuroscience, I argue that the easy and hard problems are actually inseparable. A functionalist description of a particular cognitive process is turned into a hard problem the moment one starts to ask how the individuality of the symbols used in the description is established. The marvel of the human cognition is not that we are able to have a subjective experience as separate from the functional operations in the sensorimotor domain, but rather that we are able to conduct functional operations while keeping the ability to address and to be aware of the phenomenal and "hard" aspects of subjective experience that accompany these operations when necessary. Thus, perceiving qualities of subjective experience is a part of the human meta-cognitive abilities, and is related to such cognitive capabilities as pretend play, theory of mind, and language comprehension. I argue that such a viewpoint is essential in clarifying the often confusing arguments surrounding the possibility of a computer simulation of human cognitive process, which are based on a strict distinction between what is computable and non-computable. Specifically, I argue that cognitive processes involving phenomenal qualities falls within the domain of "FAPP" (for all practical purposes) computability. I discuss how it is possible for a problem to be in the "computable" or "functionalist" domain but also "hard.

Introduction

The physical description of the world might seem to provide a causally closed theory of how the things in the world evolve, and the hope for someday attaining a "theory of everything" in this sense is still very much alive (e.g. Taubes 1995). In fact, there is up to today no grave demonstrated flaw in the physicalist's program of providing a complete theory of the objective behavior of materials in the world. The physicalist way of looking at the world is equivalent to the computational worldview, where every physical process in the world is a manifestation of some computational process. We can alternatively view the physical phenomena in the world as obeying some physical laws described by a set of equations or a result of some computational process, e.g. in the form of cellular automata. In describing the cognitive processes that occur in the brain, while basically assuming physicalism/computationalism, one may add the functionalist's point of view, the view that a description of the functionalities that the physically or computationally implemented activities of neurons have is a complete characterization of our cognitive process.

The golden triangle of physicalism, computationalism, and functionalism seems to provide a complete framework for describing everything under the sun, including our cognition. The difficulty in coming to a satisfactory explanation of the nature of the subjective experience, however, is an entirely separate issue.

Chalmers (1995) pointed out that the nature of subjective experiences is the only hard problem to explain, whereas functional accounts of cognitive processes, such as the ability to categorize, to direct one's attention to something, to remember the things one experiences in a coherent way, whether conscious or unconscious, are easy problems. Chalmers correctly pointed out that a conventional neurobiological approach, even at their best, such as the one adapted by Crick and Koch (1995), does not really address the hard problem.

Here, while basically adapting Chalmer's viewpoint, I present an alternative analysis of the problem. Specifically, I argue that the so-called easy and hard problems are actually inherently intertwined. The hard problem cannot really be "solved" or "grasped" as long as we consider it as something separate from or outside of the functionalist domain. In fact, the hard problem and the functionalist agenda go hand in hand together.

As an example, one could take an easy approach or hard one to the problem of the meaning of a word. A philosopher can argue for many hours what the word "time" means. When making an appointment with a friend, however, the same philosopher would advisedly take an "easy" approach to the meaning of time, adapting a functionalist's interpretation of such a phrase as "let's meet at the station one hour later". If the philosopher adapts a "hard" approach to the concept of time, he is going to miss the appointment.

The same applies to the problem of qualia (Jackson 1982, Chalmers 1995, Mogi 1999b). One could take the origin and nature of "redness of red" as a hard problem.

Indeed, how the qualia in our conscious experience arise as a result of the neural activities in the brain is the typical "hard problem", one that cannot be handled with the concept of "response selectivity" (Mogi 1999a), a central dogma in neuropsychology for many years. Despite that, one could equally treat the "redness of red" within the framework of an easy problem, simply pushing the button when one sees something red, when instructed to do so.

Taking these examples at hand, if one examines how the human cognitive processes occur, one is led to the conclusion that the easy and hard problems are not so separate after all. The gist of the human cognitive process is not that it is immersed entirely in the hard problem, but that we can take alternatively hard or easy approaches to our own cognition based on the "aspectual shape" (Searle 1992) relevant to the situation. We are actually able to dynamically switch between taking the phenomenal aspects of the conscious experience at their face value and just treating it in the functionalist way. The ability of treating the same process as easy and hand problems is actually the essence of human cognition.

Metacognition

The human ability to treat the same problem (such as the meaning of the concept "time" or the "redness" of red) either as an easy or a hard problem is related to the ability of metacognition, the ability to assess one's own mind state from a meta-level. When processing the command "press button when you see something red", one may not necessarily be aware of the "redness" of the object that has given the "go" signal. To be "aware" of the sensory qualities of the percepts that support one's ability to perform a certain function is an instance of metacognition, in that it is something that is added on top of the rudimental functionality of performing the right thing at the right signal. When we are aware of seeing something "red" in the visual field, we are in effect having a metacoginition of our own internal state, making a phenomenological "measurement" of the neural activity in the brain at that psychological moment.

In the above sense, metacognition is ubiquitous in our subjective experience. Whenever we are aware of the sensory qualities (qualia) that make up the phenomenal experience, we are having a kind of metacognition. In that everything that we experience consciously has a certain phenomenal quality, consciousness is in its most general sense a metacognition of one's internal state in terms of qualia. To regard a particular cognitive process as a hard problem is then dependent upon having a metacognition of phenomenal qualities accompanying the cognitive process in question. The easy approach to cognitive function is one without the metacognitive element.

In order to grasp the relation between having the capability of metacognition (and therefore being aware of having qualia in one's conscious experience) and the conventional functionalist description of cognitive processes, it is necessary to realize that metacognition is probably not a solely human phenomenon.

The ability of metacognition is likely to have emerged gradually in the history of evolution. Rhesus monkeys are able to monitor how certain one is about a memory task that one is about to perform (Hampton 2001). When given an option to avoid the task, the probability of the monkey to avoid the task appears to increase as the probability of correct answer decreases. Thus, the monkey seems to be able to assess its internal state, monitoring how certain it is about the memory task.

Being able to explicitly reflect on one's own internal state is not likely to be highly developed in lower animals. However, the rudimentary ability to be aware of the visual qualia in one's visual field, in itself a metacognitive process in that it adds something on top of the purely functional description of the visual process, is likely to have emerged before the advent of more advanced abilities to assess one's internal state within the ultimately linguistic intentional framework.

The binding of sensory features, in particular that of "property binding" (Treisman 1996) apparently occurs in terms of the various visual qualia involved (color, surface texture, transparency, motion, etc) being organized into a particular locus in the visual field. The fact that the ability to perceive the environment in terms of bound features is ubiquitous from lower animals is consistent with the assumption that the ability to have a metacognition of phenomenal sensory qualities has occurred early in evolution.

By the very nature of conscious subjective experience (we cannot in principle certify if other humans are in fact "zombies" (Chalmers 1996) or not), it is not clear if the rhesus monkey has the same kind of conscious experience associated with metacognition in humans. We do not know what it is like be a bat (Nagel 1974), but that should not in itself exclude the possibility that bats have metacognitive abilities. The difficulty of assigning metacognition to animals comes from the fact that at present it is not clear what essential new element having a metacognition of something adds on top of the conventional functionalist paradigm of the cognitive process.

Language and metacognition

It is illuminating here to consider the role the metacognition of the semantics of language plays in our daily conversation.

Metacognition is a particularly central and important issue when it comes to the semantics of language. It is difficult to pin down the meaning of any single word when you come to full grips with the semantic richness of the language. For example, the usage of the word "light" in the English language cannot be simply grasped by stating that the word "light" is a shorthand for the electromagnetic wave that one gets as a solution for the Maxwell's equation. Phrases like "she is the light in my life" certainly demonstrates the almost infinite nuances in which this particular word can be employed. The definition of the meaning of and accounting for the ultimate origin of a single word is as hard a problem as one can get in cognitive science. In this sense, the metacognition of a meaning of a word is never complete.

However, in daily life we use the words as if we understood their meaning. Taking a behaviorist's standpoint, understanding the meaning of words might be taken to imply that one acts as if one understands the meaning of the word, a philosophically practical viewpoint taken by the classic paper by Turing (1950).

To understand language in a broader context of cognitive processes, one would probably have to extend the linguistic behaviorism introduced by Turing to a kind of "neural behaviorism". The ability to understand a particular system of language is to be found in the cortical circuit of handling semantically rich intentional aspects of percepts, and does not necessarily have to be expressed explicitly as a verbal output. One does not have to be able to utter a single word to behave as if one understood the meaning of a word in a simple sense. Even pre-linguistic children, when tested by the paradigm of preferential looking, can exhibit behavior as if they understood the meaning of a word. For example, if the children are presented with a pair of pictures, one of which is a picture of an apple in various forms and the other is picture of other objects, in conjunction with the auditory stimulus of "apple", they preferentially look at a the picture of an apple longer than at the other one (Sudo and Mogi 2003).

Needless to say, the infants are not likely to have a metacognitive ability to consciously reflect on the meaning of a word. The development of the fluency of language and the ability to verbally improvise seems to coincide with the ability to reflect in a metacognitive sense on the meaning of a word, if necessary.

Taking the Turing test paradigm and the idea of a universal Turing machine that emulates human linguistic abilities at the face value, it is not at first glance apparent how the metacognitive process or the ability of addressing the hard problem helps in passing the Turing test or implementing computational abilities in a universal Turing machine.

The Computability of Cognitive processes

Faced with the apparent hard problems concerning the subjective experience, and the difficulty of grounding the phenomenal aspects of cognition on a computational or functionalist basis, some authors have referred to concepts that are generally considered to be outside the reach of physical/functional framework that is usually assumed in describing the cognitive processes in the brain as a possible remedy to the problem.

One idea that has been occasionally cited as a source of remedy to the hard problem is that of computability. Penrose (1989) claimed that an essential ingredient of human conscious cognitive process, i.e. understanding, is something that falls outside what is computable, citing the incompleteness theorem of Goedel (1931). Penrose also claimed that the non-computable aspects of human cognition could be only understood within the framework of a theory of quantum gravity, which will be presumably formulated in a non-computable framework.

As Penrose admits, at present it is difficult to come to an actual description of a non-computable process, all the more so to give a description of physiological realistic non-computable process that actually goes on in the brain. One example of non-computable process that Penrose gives involves the judgment of the halting or otherwise of a particular Turing machine. But it is difficult to envision how such a process could be implemented in the brain.

Regardless of the practical implementability of a non-computable process, it is certainly interesting and legitimate to ask whether conscious cognitive processes falls within the computable class. In particular, it is interesting to ask whether aspects of the hard problem such as qualia or the semantics of words, and our metacognitive ability to reflect on those phenomenal properties, fall within what is computable.

FAPP computability

Here, it is helpful to introduce the concept of "for all practical purposes" (FAPP). The term "for all practical purposes" was coined to provide an excuse for the standard interpretation (i.e., the Copenhagen interpretation) of the wave function reduction process in quantum mechanics (e.g. Bell 1989). There are both epistemological and ontological difficulties concerned with the wave-particle duality and wave function reduction in the Copenhagen interpretation of quantum mechanics, illustrated by the famous paradox of Schrodinger's cat. However, for all practical purposes, i.e., in terms of giving an accurate prediction for the behavior of particles within the limits of knowability set

by the general principles of quantum mechanics, the Copenhagen interpretation seems to work without any practical flaw. The Copenhagen version of quantum mechanics is a complete theory FAPP, although gravely unsatisfactory from the point of view of epistemologically complete description of the world.

There's an analogy between the point of view that the Copenhagen version of quantum mechanics is complete and the idea that it is possible to simulate all objective aspects of human cognitive processes by computational means. One could extend behaviorism to neural behaviorism, and claim that the objective behavior of neurons, whether it is explicitly reflected on external bodily movements or not, gives a complete description of human cognitive process. Given the universality of Turing machines, in addiction to the general freedom involved in establishing a mapping between the behavior of physical systems (even those as complex as the human brain), it is difficult to envisage an instance of knowably non-computable process that is representative of human cognition. Our intuition that the phenomenal aspects of our cognition cannot be explained by any computational models is a strong one. Therein lies the ultimate source of inspiration for arguments like the one by Penrose. However, it is difficult to envisage an aspect of the cognitive process that is knowably non-computational.

In the rigorous mathematical sense, the difference between what is computable and what is not is a clear one. On the other hand, in considering the mapping between numerical reality and the physical reality, it is not always apparent if the watershed between what is computable and what is not is relevant in physical reality. One cannot compute an irrational number by means of a finite process in a Turing machine with finite number of states. When irrational numbers are non-computable in a system, and only rational numbers can be computed, whether a state is to be represented by a rational or irrational number makes a huge difference in mathematical reality. However, in real physical situations, it is not apparent whether the distinction between a rational and irrational number is relevant or not. One can find infinitesimally close rational number (which one can compute) given any irrational number (which one may not be able to compute). In mapping the real physical processes (on which the conscious cognitive process ultimately supervene) to numbers, the difference between a rational number and irrational number does not appear to make any practical difference.

In a deep philosophical sense, it is not apparent how numbers are materialized in the physical process in the world. The relevance of computability is heavily dependent on the relation between the numerical and the physical. In arguing about computability, the difference between rational and irrational numbers makes a whole world of difference. In the real world, we are very much hard pressed to find any process cognitive or otherwise which critically hinges upon a distinction between what is computable and what is not.

For all practical purposes, therefore, it seems reasonable to assume that all cognitive processes, whether conscious or unconscious, are computable. The FAPP computability of cognitive processes, including those conscious processes involving qualia and self-consciousness, has important implications for the nature of conscious experience, and its relation to such conventional viewpoints as functionalism and connectionism.

Consciousness and functionalism

If we embrace the FAPP computability of conscious process, the arguments on consciousness dependent on the hypothesized non-computability of conscious process become irrelevant.

The Weltanschauung that needs to be taken, then, is one where the phenomenological and functionalist descriptions of the cognitive processes are intertwined. The phenomenal aspect of subjective experience is not something that is beyond or outside functionalism, but something that goes with it.

When we reflect on the way that the human cognitive processes are actually organized, we discover that what is on the surface within the domain of functionalism is tightly coupled with the phenomenal properties of percepts, especially when metacognition is involved. When performing a simple numerical manipulation such as Euclid's algorithm, we may appear to be performing a purely computational task. When metacognition of the sensory qualities (qualia) of the percepts involved in the apparently computational process is taken into consideration, however, we discover that qualia are involved in, and underlie, the computational operations *as they are perceived* in our subjective experience. The easy and hard aspects of our cognitive processes are thus tightly intertwined, while everything can be grasped in the FAPP computable domain as far as the objective behavior is concerned.

The significance of the human cognitive abilities to switch between the easy and hard stances is clear in such cases as pretend play and theory of mind (Baron-Cohen et al. 1985). It is in general a hard problem to be able to pretend to be somebody else, or to come to an understanding of other's mind. Taking an "easy" approach, one may perform cognitive tasks starting from a simple assumption, e.g. in such cases as a boy pretending to be a girl or guessing what caused the sadness of one's friend based on external observation and knowledge. One may even claim that some aspects of the pretend play and theory of mind can be implemented as an algorithm. On the other hand, when

coming to terms to the phenomenal aspects of actual subjective experience of being a girl or one's sad friend, hard aspects of the problem suddenly emerge. In our daily cognitive activities, we basically perform the tasks adapting the easy stance, while reflecting on the metacognitive aspects that eventually lead to the hard problem when relevant. The fact that the human being can lead a highly social life is very much dependent on the ability to treat the potentially hard problem of pretending and guessing other's mind as an easy problem, while maintaining the metacognitive ability to immerse oneself in the hard problem when necessary.

Conclusions

Nobody can deny that conscious experience exists.

I once had an argument with a colleague of mine in University of Cambridge, U.K., when he claimed that there is no such a thing as consciousness. When I faced him with the question "but surely you admit that there is a clear difference in your subjective experience between when you are awake and when you are sleeping", he claimed that as far as he was concerned, there was no essential subjective difference between the awake state and the sleeping state (Pelah, 1997). I am still hard pressed to come to a reasonable interpretation of his remark, although it was certainly a consistent one from somebody who denied the existence of consciousness.

The idea that the functional or computational description of the neural processes in the brain is all that is required is epistemologically unsatisfactory but probably FAPP true. The great difficulty experienced by those trying to go beyond the functionalist kingdom is probably due to the fact that people have been looking at the wrong places.

The reality of the human cognitive processes is that the easy and hard problems are inseparable. Our ability to switch between the easy and hard stance is the foundation for the flexible adaptation to the natural and social environment. The solution for the admittedly difficult problem of accounting for the phenomenal qualities of our subjective experience is to be sought in the delicate manner in which the physicalist/computationalist/functionalist description can go hand in hand in our cognition. The neural mechanism underlying metacognition, in the broadest sense of the term including awareness of sensory qualia, would provide the essential springboard for the understanding of consciousness.

Acknowledgments

I thank my colleagues Takashi Ikegami and Yukio Gunji for helpful discussions concerning the issues addressed in this paper.

References

Baron-Cohen, S., Leslie,, A., Frith, U. (1985) Does the autistic child have theory of mind? Cognition 21, 37-46.

Bell, J.S. (1989) The speakable and unspeakable in quantum mechanics. Cambridge University Press

Chalmers, D. (1995) Facing up to the problem of consciousness. Journal of Consciousness Studues. 2(3), 200-219.Chalmers, D. (1996) The Conscious Mind. Oxford University Press.

Chalmers, D. (1996) The Conscious Mind. Oxford University Press.

Crick, F. and Koch, K.Crick F, Koch C (1995) Are we aware of neural activity in primary visual cortex? *Nature* 375:121-123.

Goedel, K., (1931) "Ueber formal unentscheidbare Saetze. der Principia Mathematica und verwandter Systeme I",. Monatsh. Math. Phys. 38 pp. 173-198

Hampton, R.R. (2001) Rhesus monkeys know when they remember PNAS 98, 5359-5362

Jackson, F. (1982) Epiphenomenal Qualia. Philosophical Quarterly, 32, 127, 127-136

Mogi, K. (1999a) Response Selectivity, Neuron Doctrine, and Mach's Principle. in Riegler, A. & Peschl, M. (eds.) Understanding Representation in the Cognitive Sciences. New York: Plenum Press. 127-134.

Mogi, K. (1999b) Supervenience and qualia. Behavioural and Brain Sciences. 22(5). pp.84-85

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review* 4:435-50.

Pelah, A. (1997) Private communication.

Penrose, R. (1989) The Emperor's New Mind. Oxford University Press.

Searle, J. (1992) The rediscovery of mind. MIT Press.

Sudo, T. and Mogi,. K.(2003) Pre-linguistic category formation in infants. Proceedings of the ICCS 2003, in press.

Taubes, G. (1995) A Theory of Everything Takes Shape. Science, 269, 1511-1513.

Treisman, A. (1996). The binding problem. Current Opinion in Neurobiology, 6, 171-178.

Turing, A. (1950) Computing Machinery and Intelligence, Mind 49, pp 433-460