The Essential Role of Consciousness in Mathematical Cognition

Robert F. Hadley (Hadley@CS.Sfu.ca)
School of Computing Science, Simon Fraser University
8888 University Drive, Burnaby, B.C., V5A 1S6 Canada

Abstract

In his most comprehensive book on the subject (1994), Roger Penrose provides arguments to demonstrate that there are aspects of human understanding which could not, in principle, be attained by any purely computational system. His central argument relies crucially on renowned theorems proven by Gödel and Turing. However, that key argument has been the subject of numerous trenchant critiques, which is unfortunate if one believes Penrose's conclusions to be plausible. In the present article, alternative arguments are offered in support of Penrose-like conclusions (although the present arguments differ markedly from his). It is argued here that a purely computational agent, which lacked conscious awareness, would be incapable of possessing crucial concepts and of understanding certain kinds of geometrically-based proofs.

Keywords: Consciousness; Cognition; Penrose; Infinity;

1. Introduction

In his most comprehensive book on the subject (1994), Roger Penrose provides arguments to demonstrate that there are aspects of human understanding which could not, in principle, be attained by any purely computational system. In particular, Penrose argues that human mathematicians are capable of proving propositions (specifically, certain metatheorems) that could not be proven by any computer program. Penrose argues further that this superiority of humans over machines (including the most advanced kinds of AI systems), within specific realms of mathematics, also entails that computers are incapable of understanding the semantics of mathematical formulae. The crux of Penrose's case involves his appeal to renowned theorems proven by Gödel and by Turing. These theorems establish limits upon what can be derived within formal deductive systems devoted to arithmetic, and upon what can be discovered by any computer program that accepts computer programs as input. We need not explore the details of those theorems here, but interested readers will find a concise discussion of those results in my recent paper (Hadley, 2008). For present purposes, we need only note that numerous researchers have challenged the case that Penrose has constructed, in both of his books on this topic (Penrose, 1989, 1994). While everyone accepts the correctness of the theorems established by Gödel and Turing, many commentators have argued, for a variety of reasons, that Penrose's application of those theorems within the cognitive realm is seriously flawed. A very common criticism concerns his contention that if mathematicians are following purely computational methods when they discover theorems, then those methods would be logically sound, and any correct formal model of those methods would be logically consistent (Feferman, 1995; McCullough, 1995). I will spare the reader most details, but my own critique, given in (Hadley, 2008), presents a rigourous argument demonstrating that even when an AI system employs only correct algorithms, and makes only valid inferences, an inconsistency may result when that system is directed at an accurate formal deductive model of itself. If such an inconsistency could result when an automated agent (AI system) is employed, it could likewise result in the case of the best human mathematicians. Another major difficulty is that Penrose appears to assume (mistakenly) that if human mathematicians are indeed following reasonable computational processes, in their search for proofs, then they must be employing infallible algorithms, rather than some very enlightened, but fallible search process.

Despite doubts regarding crucial aspects of Penrose's arguments, I nevertheless believe that some of his conclusions are both correct and important. In particular, I agree that a purely computational agent, which lacked awareness or consciousness, would also lack the capacity to understand some crucial mathematical concepts and forms of reasoning. Indeed, I will argue in support of this general conclusion. I will do so by addressing the problematic nature (computationally speaking) of the concepts of countably infinite sets and of non-denumerable (higher-order) infinities. Arguments presented there differ markedly from those due to Penrose. Following that, I argue in favour of a suggestion of Penrose (unproven hitherto) regarding the inability of purely computational agents to grasp the reasoning found in crucial, geometrically-based proofs (Note, the awareness required to grasp such concepts might be present in a computational agent, but only as a contingent side-effect of a variety of factors. I return to this issue.) This latter section evokes to some degree certain themes found in Searle (1992). However, the specific arguments I present are not Searle's. (It is noteworthy that many researchers in Artificial Intelligence remain unconvinced by Searle's famous 'Chinese Room' argument, and appear quite willing to attribute understanding to a hypothetical robot which could merely pass certain behavioural criteria. It should be valuable, therefore, to consider arguments which specifically support Penrose's conclusions, but which do not depend upon an acceptance of Searle's famous argument.)

Before delving into details of the arguments presented below, it would be well to pause briefly, to consider the sense of 'consciousness' and 'awareness' at issue here. Although I will not attempt to define these terms (and do not think definitions are actually possible for the concepts we commonly associate with the words), a few illustrative remarks may be helpful. In using these terms, I have in mind consciousness (or awareness) that is focused on a specific topic. For example, we can become conscious of having particular thoughts (about things as various as food, pains, tactics, and sets of numbers). Also, we can become conscious or aware that we perceive some geometric figure or that we understand certain concepts (such as the concept of the set of natural numbers).

Moreover, very often we are consciously thinking or perceiving on specific occasions, without being conscious or aware that we are thinking or perceiving. There may be other subtle differences among the kinds I consciousness I have mentioned here, but for present purposes I will not attempt an analysis. Indeed, it is noteworthy that if operational or behavioural analyses were possible in these instances, the famous problem of other minds would arguably not exist.

2. Human Concepts of Infinite Sets

Let us now consider whether, in all cases, the semantic understanding of human mathematicians can be adequately captured by fully computable, operational procedures. Consider first the concept of a countably infinite (denumerable) set of objects. Typically, students of mathematics have acquired a reasonably good understanding of this concept by the end of their first course in differential calculus, if not earlier. The concept is frequently explained to students by an informal definition along the following lines: 'a countable (or denumerable) infinity of objects is an infinite set of items, where each element of the set can be put into one-to-one correspondence with a distinct natural number.'[1]

It is crucial to note, however, that this 'definition' already presupposes an understanding of an infinite set, in two separate ways. First, the definition explicitly appeals to the meaning of the word 'infinite'; secondly, the definition invokes the concept of the natural numbers, which, of course, is again a concept of an infinite set of elements. It would appear, then, that any attempt to provide a robot (or other AI system) with the concept of a countably infinite set, by providing the robot with the definition just considered, would involve a serious circularity. It might be suggested that the circularity could be diminished, at least, by replacing the word 'infinite', in the quoted sentence, with the phrase 'extremely large'. Such a replacement raises other difficulties, but let that pass. For, even to assume that the robot could understand the concept of the set of natural numbers is already to assume that it has grasped the concept of a countable infinite set. (Indeed, a pleasing property of the natural numbers is that they are paradigmatically a countable set – they come numbered!) What is worse, from the standpoint of those who seek to reduce all human thought to computational processes, is that every attempt to define, explicitly, the concept of a countable infinity must inevitably invoke the concept of the natural numbers.

However, let us consider whether we could, by purely computational means, endow a robot with the concept of the set of natural numbers. For, if we could do that much, we may have gone far towards endowing it with the notion of a denumerable infinity. This much is clear; we can certainly provide a robot with a (non-terminating) procedure for generating the Arabic numerals which individually each express one of the natural numbers. There is a simple non-terminating procedure for generating, say, each of the decimal or binary numerals. Of course, at no time will a computer or a human ever have generated the entire set of numerals, so, it would be absurd to suggest that a robot (or a human) would grasp the concept of the entire set of natural numbers by having produced the set. It might be suggested, though, that we could 'tell' the robot that the process of generating the 'entire set' will not (or could not) halt. For argument's sake, let us grant that the robot might already possess the concept of negation (the sense of 'not') and the concept of halting. We may further assume then, that, via ordinary semantic compositionality, the robot could come to understand the sense of 'not halting'. By similar charitable assumptions, let us concede that the robot understands the meaning of 'the process never halts'. In this fashion, we might allow that our robot could, in solely computational terms, come to 'understand' that it is always possible to generate one more numeral, or even that 'no matter how many numerals have so far been generated, this robot can generate at least one more'.

On the face of it, it might now appear that our robot has thus been endowed with a complete, human-like conception of the infinite set of countable numerals. (To be sure, numerals are merely names and are not themselves numbers. So, our robot would not yet have acquired a genuine conception of the set of natural numbers, but let that pass.) However, we should resist this superficial appearance, as I shall now argue. A crucial aspect of many mathematicians' understanding of even a countable (enumerable) set of items (in the case numerals) is that the entire set of elements can (or does) simultaneously exist. The human concept of this infinite set is not just that of a finite set, so far generated, combined with the potential for always generating yet another element of the set. Rather, most mathematicians (and even many students of mathematics) conceive and realize that, over and above the elements already generated or enumerated, there exist vastly many more elements that can be enumerated. Indeed, not only do they realize that vastly many elements remain to be enumerated, but infinitely many elements yet remain. The italicized phrase here is crucial. We cannot hope to endow a robot with a purely computational understanding of a countable infinity if part of what we must make the robot understand is that, no matter how many items it has generated or enumerated, there remain infinitely more items to come. Moreover, the robot would need to understand that all these remaining items are already contained in the infinite set in question.

It is puzzling, to be sure, how a human student could come to grasp the seemingly circular aspects which I have just stressed. It may well be that, initially, the student's understanding of infinity does merely consist of notions

[1] The set of natural numbers includes zero, together with all the positive integers (0, 1, 2, 3 ...).

such as ‘there is no end to this process of generating numerals’, and ‘I can always keep generating one more item’. However, in the case of many intelligent students at least, there comes a stage where the student’s understanding leaps to a higher level. The student’s awareness of the items being enumerated, together with the unending nature of the enumerative (or generative) process, enables a conceptual or imaginative leap to occur, and in this process the student is able to entertain the idea that the entire aggregate set of elements simultaneously exists. Of course, the student does not visualize the entire set, but may, in some schematic fashion, imagine a series of elements stretching off into a vanishing point. In any case, I submit that the conceptual leap that occurs involves an emergent awareness, and at the very least, significant meta-processing must be involved.

Please note, however, that I am not claiming that emergent awareness and metaprocessing could never occur within a computational robot. Rather, my contention is that such emergence and meta-processing are not to be found in any given computational procedure which purports specifically to express or embody the concept of a ‘countable infinity’. If it someday happens that we discover computational processes which could engender emergent awareness of a type sufficient to enable the ‘conceptual leap’ mentioned in the previous paragraph, then these processes would very probably enable a broad range of conceptual leaps. That is, the engendered awareness would be of a fairly general nature. I say this because the awareness which presently enables us to imagine an unbounded set of objects, simultaneously co-existing, does not in any way seem to be specific to the particular concept of an infinite set. (We should bear in mind, also, that for all we know, consciousness can emerge only as a combined side-effect of both chemical processes and specialized neurological events. Moreover, and crucially, computer simulations of chemical processes do not produce the same effects as the chemical processes themselves (as Searle has stressed). Consider a computer simulation of hydrochloric acid dissolving iron.)

Perhaps, however, it will now be objected that there are eminent mathematicians who belong to the intuitionist, or constructivist school of thought. Members of this ‘school’ commonly deny the existence of any complete, actualized set of infinite objects, such as the set of natural numbers. It may appear that these mathematicians, at least, cannot conceive of a fully realized countable infinity, and do not possess the concept that I have been arguing for.

In reply, I would stress two points. In the first place, mathematical intuitionists (or constructivists) do not claim to be unable to conceive of a completed, fully realized infinity. Rather, they simply doubt or deny the existence of the set of objects. Put another way, they doubt the existence of the extension of the phrase, ‘countably infinite set’, rather than the intension of the phrase.

Secondly, even if some intuitionists insist that they cannot understand the meaning of ‘countably infinite set’, or claim to lack the corresponding concept, there are certainly many other mathematicians, belonging to the Platonist school of thought, who are certain they do understand the concepts, not only of countable infinite sets, but of larger, nondenumerable infinite sets. Surely, the cognitive abilities of these Platonically inclined mathematicians fall within the scope of Cognitive Science as much as the mathematical intuitionists do.

Having now considered reasons to believe that the concept of a countable infinity is not expressible in terms of a computational procedure, let us turn to conceptions of higher-order, non-denumerable infinities. Many, though not all, mathematicians have accepted the work of the renowned Georg Cantor, who developed (via subtle proofs) a theory of transfinite numbers, based upon a hierarchy of infinite sets. The first (or lowest order) of these infinities are the countably infinite sets that we have already considered. All countably infinite sets are judged to be of the same magnitude (i.e., cardinality), and that magnitude was termed by Cantor, aleph-null. Immediately following the lowest order infinity, is the class of infinite sets whose magnitude (cardinality) equals that of the set of real numbers (the latter includes all the transcendental numbers (e.g., □), all ‘rational’ numbers, and all integers). The cardinality of the set of real numbers constitutes a 2nd-order infinity, known as aleph-one. The size of aleph-one vastly exceeds that of aleph-null, but one can begin to get a glimmering of the magnitude involved by considering that any given countable infinity can be mapped to an arbitrarily small segment of the real number continuum (which we assume extends infinitely far to the left and right). We can, therefore, map an infinity of countable infinities into the entire real continuum, whose magnitude is aleph-one. Without going into further technicalities, the crucial point to note is that one cannot acquire the concept of aleph-one, the first non-denumerable infinity, unless one already possesses the concept of a countably infinite set (whose magnitude is aleph-null). Indeed, each of the higher order, non-denumerable infinities is defined in terms of the preceding order infinity, so that the semantics of each of the transfinite ‘numerals’ (aleph-one, aleph-two, etc.) ultimately presupposes the concept of the countable infinity, aleph-null. It follows, then, that if my preceding arguments are accepted, and the concept of a countable infinity cannot be captured by any computational procedure specific to that concept, then the same limitation applies to the concepts of each of the higher order infinities. (Note, by the way, that none of the higher-order infinities, beginning with aleph-one, could remotely be defined in terms of such phrases as ‘the set of items is unbounded’, or ‘however many items you have counted, there will always be more to count’. Such expressions do not begin to convey the magnitude even of aleph-one, which corresponds to the cardinality of the set of real numbers.)

3. Human Perception of Geometric Diagrams

Apart from the human capacity to form conceptions of a variety of types of infinity, there is a crucial human ability, involved in our capacity to understand certain geometrically based proofs, which bears upon our overall concerns here. Specifically, there is strong reason to believe that the relevant kind of understanding necessarily involves conscious awareness in the perception of complex geometric diagrams. This perceptual awareness is intimately related to what cognitive psychologists identify as the 'perception of a gestalt'. Bearing in mind that, for all we know, conscious awareness is not a mere by-product of computational processes, but may well be partially the result of complex chemical 12 processes, the following arguments, should lend credence to Penrose's assertion that the ability to understand certain mathematical arguments cannot merely be the result of the execution of computer algorithms. (N.B. some of the following points are evocative of the work of John Searle, 1992, but I have considerably reduced the number of assumptions required, and my conclusions are much less sweeping than those of Searle.) Let us note, first of all, that the ability to perceive certain schematic diagrams as instantiations of geometric shapes appears, crucially, to require awareness. For example, consider the seven 'dots' displayed in Figure 1.

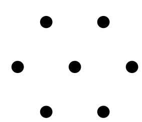

Figure 1: Seven 'dots' which schematically represent a hexagon.

Very little effort is required, typically, for an adult human to perceive these seven dots as comprising a hexagon. (Indeed, considerably greater effort is required to simply perceive the collection of dots merely as a set of unrelated dots.) Moreover, humans perceive the hexagon shape virtually instantly. Certainly, no conscious analysis or dissection of the 'data' is involved. Let us consider, however, whether a purely computational process could come to perceive the collection of dots as a hexagon. There are two possible cases to consider.

In Case 1, a sequential computer program is presented with a digitized image of the 'set of dots', and from this image the program is able to draw lines between dots lying on the perimeter, compute angles between lines that share common vertices, and engage in other computational processes. Without doubt, an appropriately designed program would eventually identify the set of dots as comprising a hexagon, and would label it as such. However, there is nothing about such a computational procedure which should tempt us to suppose that the computer 'perceives' the set of dots as a unified object. The mere algorithmic processing of a series of dots, involving the construction of line segments, measurement of angles, etc., in no way logically entails that any postulated awareness of the separate discrete elements involved (dots, lines, angles ...) would result in an awareness of a single, cohesive object. Yet, anything that is purely the result of a computational process would be logically entailed by formal analysis of that same process. (See Hadley, 2008; Penrose, 1994; or Kleene, 1967, for an explanation of the mathematical relationship between computable procedures and logical entailment.) It follows that if any computable, sequential procedure does manage to perceive the collection of seven dots in Figure 1 as a coherent, unified hexagon, this perception does not result as a computationally necessitated consequence of the computer program that instantiates the algorithm involved. Rather, any such unified perception would almost certainly be, at best, an emergent, contingent side-effect of the program's execution. Consider then Case 2, according to which a computer program comes to identify, purely via parallel algorithmic processing, Figure 1 as an instance of a hexagon. In the case of such parallel processing, the seven separate dots in figure 1 would be simultaneously represented and processed by concurrently active 'memory units' and procedures within the computer. Would this simultaneous representation and processing of seven, separate discrete dots necessarily result in the perception of a cohesive, unified hexagon (assuming, as before, that an appropriately designed program was involved)?

The answer, once again, is no. For the mere fact that memory units, representing the seven separate dots, are simultaneously activated, or even that line segments between the adjacent dots lying on the perimeter are concurrently 'drawn' would not automatically result in a realization that all these dots and line segments belong together. Even if we assume that some awareness is present in the execution of the parallel procedures, there is no logical reason why that awareness should not be separately, albeit concurrently, focused on the separate dots and line segments involved. There is no more reason to suppose that a unified perception would result than there is to suppose that six different people, each located several miles from one another and each simultaneously examining a single large dot which has been connected by a visible line to some corresponding dot that is likewise viewed by a 'fellow examiner', would result in a unified perception of a very large hexagon. Even if we suppose that the people involved can each talk to one another via cell phones, and that they collectively come to deduce that some hexagon must be present, there is still no logical compulsion to suppose that the separate people collectively perceive a single hexagon.

Note, moreover, that it does not help matters to shrink the scale involved and to replace the separate people by collections of separate neurons in a human brain. For, the fundamental problem still remains; the separate neurons, even though they may send signals to one another (just as the people were able to talk via cell phones) do not share a single awareness in which a perception could be 'unified' (indeed, the separate neurons presumably have no awareness at all). It may well be that the simultaneous activation of separate collections of neurons (which are concurrently stimulated by the seven distinct dots in Figure 1) would, in fact, when coupled with appropriate chemical reactions and the exchange of electrical signals over a suitably configured network, result in a unified perception of a hexagon. However, this result would be causally produced, and not be logically entailed. There is no way that we can deduce, logically that a unified perception would result in this case. Given this, we are assured that no computer program, whether parallel in design or not, could engender the required unified perception of a hexagon purely by virtue of the execution of the program itself. If a unified perception results, it is an emergent by-product of a non-logical order.

In light of the foregoing conclusions, we are in a position to see that the human capacity to follow (and indeed to discover) certain geometrically-based proofs requires the presence of conscious, gestalt-like perceptions. A fragment from one of Penrose's geometric proofs will be helpful here (although Penrose uses the proof for a very different purpose). As Penrose explains (Penrose, 1994) there is an intriguing, provable relationship between those natural numbers known as 'hexagonal numbers', and those which are perfect cubes. Hexagonal numbers are so-named because their magnitude exactly corresponds to the number of 'dots' required to comprise a 'filled-in' hexagon of the kind displayed in figure 1, above. In figure 1 we see seven dots. Thus, seven is a hexagonal number. The next two hexagonal numbers, in ascending order, are 19 and 37. The hexagon corresponding to 19 can be created by the mere expedient of adding a new perimeter of dots to the hexagon corresponding to the number 7, as shown below in Figure 2.

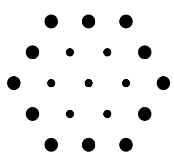

Figure 2: Enlarging a hexagon via adding a new perimeter.

In a similar fashion, we could add a perimeter of dots to Figure 2 to obtain a hexagon corresponding to the next hexagonal number, which is 37. The complete set of hexagonal numbers is infinite, and for each such number one could, in principle, generate a corresponding geometric figure in the fashion just illustrated. Now, it turns out that if we take the number '1' to be hexagonal, and if we sum up any consecutive series of hexagonal numbers, beginning with '1', then the sum will be a perfect cube. (Call this last assertion, 'Theorem 1'.) For example, if we sum the first three hexagonal numbers, 1, 7, and 19, we obtain the cube, 27. By means of a graphic proof, involving 'dotted' hexagons and geometric cubes, Penrose proves the truth of Theorem 1. One step of this proof involves noting that every numerical perfect cube, N, corresponds to a geometric cube containing N units, as displayed in Figure 3.

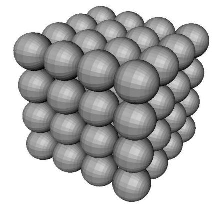

Figure 3: The volume of a geometric cube corresponds to an algebraic cube.

The next step is to note that the perimeter of any cube, when viewed from the same perspective as used in Figure 3, will be a hexagon. Significantly, humans attain this insight via a conscious perception. We see both the geometric cube and the hexagonal border my means of integrated, cohesive gestalts. Another key step in Penrose's proof involves the perception that the number of 'units' in a 'filled-in' geometric cube (such as that in figure 3) can be obtained by summing up a progressive series of fragments of cubic arrays. Such a series is displayed in Figure 4. Each fragment in the series, apart from the initial solitary unit at the bottom left, consists of a 'back wall, side wall, and ceiling', as Penrose observes (Penrose, 1994). By nesting the successive fragments together, in sequence, we can perceive with our conscious imagination that, when closely nested together, the fragments will constitute the completed, unified cube (having dimensions of 4 by 4 by 4, in the present case).

A final, crucial, step in the proof requires that we imagine viewing any given fragment (from Figure 4) in the series from a distant point opposite the vertex point of the 'three walls'. For example, if we view the rightmost fragment in the series, from this perspective, we will discover that it appears as the hexagon shown in Figure 5, below. By means of conscious, gestalt-based perceptions such as the present one, we are able to realize that the number of units in each of the fragments of figure 4, is equal to the number of units in the corresponding hexagon.

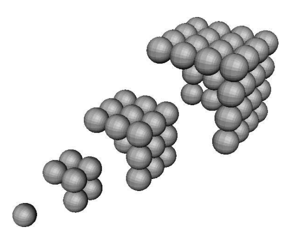

Figure 4: A series of cube fragments.

By now, it is presumably evident to the reader that the ability to understand and follow Penrose's geometrically-based proof intrinsically requires the ability to perceive complex diagrams as integrated gestalt patterns. Moreover, as I have previously argued, that ability in turn necessitates our conscious awareness. Of course, it may well be that a purely computational agent, lacking all awareness, could construct a proof of the same 'theorem' (Theorem 1). And this proof (when viewed abstractly, with many details suppressed) might possess an overall structure that bore some resemblance to the structure of the Penrose proof that we have been considering. Nevertheless, the Penrose proof is not the same proof as the one (call it Proof-2) the purely computational agent would create. For, unlike the Penrose proof, Proof-2 would lack the gestalt perceptions of hexagons and cube fragments, but would necessarily contain much detailed analysis of spatial relationships of the small constituents of the images in figures 2, 3, 4, and 5.

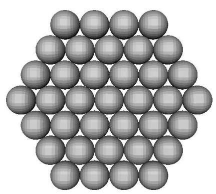

Figure 5: The view of a 4 by 4 cube, from a particular vantage point.

It might be objected, however, that when humans follow geometric proofs of the kind offered by Penrose, they are still performing, at some subconscious level, much detailed processing of small constituents, in a fashion comparable to the 'detailed analysis' which must exist in the computer-generated Proof-2. Whether or not that is so, I submit that the various, conscious gestalt perceptions employed in the Penrose-style proof are essential to our understanding the proof. It is by means of these gestalt perceptions (inter alia) that we become convinced that each step of the proof follows from the preceding one. Moreover (and I think this is of fundamental importance), the kind of gestalt perceptions we have been discussing are crucial to the discovery process for geometric proofs of the present type. At the very least, such high-level, gestalt perceptions provide powerful heuristic evidence which helps to guide the discovery process. If our brains were limited to the kind of low-level, tedious analysis that occur in proofs exemplified by Proof-2, the search space for geometric proofs would be overwhelmingly vast, and many fewer interesting proofs would be discovered.

4. Summary and Discussion

In the foregoing, I have presented arguments, grounded in mathematical domains, to demonstrate that the acquisition of human-like concepts of countable and non-denumerable infinities, and human-like comprehension of a particular geometrically motivated proof does require conscious apprehension of the subject matter involved. I have not precluded the possibility that a computational agent might come to possess the requisite consciousness, but have argued that if this consciousness does arise within the agent, it does so as an emergent, contingent side-effect of the underlying processes involved.

Moreover, I have emphasized that it is presently unknown whether the relevant consciousness could in fact arise solely as a consequence of the underlying computational processes. For all we presently know, computational processing may be, at most, just one among several causal conditions for the creation of the conscious conceivings, realizations, and perceptions involved. It may well be that special kinds of chemical activities comprise an additional necessary condition for the production of such consciousness. Alternatively, it may be that the 'quantum-classical' hypotheses advanced by Penrose (1994) are correct, and that no agent whose processing is entirely deterministic could ever be conscious. Specifically, in Part II of his scientific tour de force, Penrose offers a theory of how entirely non-computational, processes, arising at the interface between quantum and classical physics, and occurring within microtubules in the synaptic junctions of neurons, may be a causally necessary condition for the production of conscious thinking. Crucially, within Penrose's theory, the relevant non-computational processes are strongly non-deterministic; they cannot even be computationally approximated using probabilistic equations and/or random number generators.

It is essential to note, moreover, that we need not embrace Penrose's controversial, Gödel-Turing based arguments (concerning the limitations of computational provability) in order to grant the credibility of his conjectures about the genesis of consciousness.

Penrose may well be right about the nature of this genesis. If he is indeed correct, we may conclude, given the supporting arguments I have presented above, that no deterministic computer program could acquire the mathematical concepts or comprehend the mathematical proofs that I have discussed.

References

Kleene, S.C. (1967). *Mathematical Logic*. New York: John Wiley & Sons, Inc.

Feferman, S. (1995). 'Penrose's Gödelian Argument', *Psyche*, 2(7). http://psyche.cs.monash.edu.au/v2/psyche-2-07-feferman.html

Hadley, R.F. (2008). 'Consistency, Turing computability and Gödel's first incompleteness theorem'. *Minds and Machines*, 18, pp. 1-15.

McCullough, D. (1995). 'Can humans escape Gödel?'. *Psyche*, 2(4). http://psyche.cs.monash.edu.au/v2/psyche-2-04-mccullough.html

Penrose, R. (1994). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford: Oxford University Press.

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

Searle, J. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.