"Setting the scene: the claim and the issue"

"Setting the scene: the claim and the issue" is an article written by Roger Penrose which analyze the peculiarities of the human brain to answer the old question: will the computer ever achieve a genuine artificial intelligence, that is a kind of intelligence which is aware of itself and conscious of its actions? Such a kind of artificial intelligence could be compared to the human one and maybe it would become also superior. At the moment, like the Turing test shows, the computers cannot behave as humans during a normal conversation neither if it lasts only some minutes but if we analyze the rate at which computer technology has moved forward over the past half century we would start to consider the answer to the previous question not so foreseen. Penrose observes that to answer such a question we should before solve other questions like: what is the mind? How are the consciousness and the awareness generated? Actually Penrose considers consciousness and awareness two fundamental characteristics of the human mind. (1)
He presents four different possible viewpoints about the argument:

1. All thinking are computation, the feelings and the consciousness are a product of these computation. According to this theory, sometimes called strong AI, a day we will be able to create an artificial mind which could reply to all our question like a human would do. Furthermore an internal consciousness would correspond to the external manifestations of intelligence so we will be able to create a kind of robots with a kind of intelligence very similar, even superior to, the human one.

2. The brain carries out actions which generate the consciousness but although we can simulate these actions, the simulation is different from the event. The philosopher John Searle argued that a simulation of one event is different from the event itself because it is done in an electronic environment which is different from the original one.

3. The consciousness cannot be simulated computationally even if it is evoked by physical action of the brain.

4. There is no scientific way to explain the awareness. This position is known as Mysticism.

Penrose refuses completely the last position because he is confident in the power of the value of the scientific criteria. He thinks that even if the processes involved in the human awareness are not totally clear at the moment they might be explained in the future. (2) His theory is nearer to the third position; he proves that in our conscious thoughts there is a non algorithmic component which cannot be simulated by an artificial mind based on the today's concept of computers. Indeed these computers are based on computational sound algorithms, but for the Godel's theorem there are some sound mathematical propositions which are impossible to ascertain using another knowable sound algorithm. On the contrary because the human insights were able to find this kind of sound propositions it means that our understanding cannot be completely formalized as only computational algorithms. So he strongly criticizes the first position in his book Emperor's New Mind (Penrose, 1989), he thinks that it is a position with a considerable force but it is wrong.
My opinion is the same. I think that if a program could behave like a human using a huge rule book about the human behaviour, it does not mean it can understand what it is doing, like the Searle's Chinese room example shows. It will never be able to deduct the semantic of the real world from the syntax that we could teach it. Actually for the reason that I will explain deeper later on, my position is

really close to the Penrose's one although different in some ways. Furthermore I believe that our brain often does not behave in a computational way, Penrose's argument proves that for mathematical thought processes but it is also more evident in subjects like art and natural sciences that the human creative capacity is not undergone to the rigid constraints of a computer's calculations. (3)
Then in this article, Penrose examines seven alternative procedures that might be used by the human mathematicians instead of this non algorithmic element.
One of these is that we use a horrendously complicated unknowable algorithm X. (4) His argument could be split in two parts. I totally agree with the first one where Penrose applies the Godel's theorem and he proves that this algorithm X cannot be unknown and with a logical argument he explains how it is difficult that it would be horrendous complicated. Penrose starts analyzing the formal systems to verify if this unknown algorithm X could be represent by these but he notices that they cannot encompass the entirety of the insights that are available to mathematicians. Indeed if an argument is based on a formal system F, which is unknown, like X, in the meaning that its soundness is unknown, then either F is sound or it is not.
But if F is sound for the Godel's procedure we know that a sound proposition G will exist that is not possible to ascertain from F. Otherwise if F is unsound it is useless because it could yield false propositions. Since our thought processes cannot be explained only with an unknowable algorithm X, Penrose analyzes if this X could be horrendously complicated but in that case it would not be logical at all, because the mathematicians need to break down their arguments into steps which should represent things obvious to all and it would be strange that the basis of these simply steps is a horrendously complicated algorithm X.
Then Penrose introduces another reason against this algorithm X that is it does not make sense that the natural selection of the human species Evolution allowed such an X arose. Actually X do not add a significant direct survival value to our species. About this argument I disagree with Penrose because this X would not be the only exception to the natural selection mechanism. We might think about the feeling of piety: not only it is not useful to survive but on the contrary it could be a threat for the person because it leads the individual to give up his property for someone else, nevertheless it exists. Furthermore we might think about the artistic talent, it useless to our survival but it is so strong in our mind that some of us, the painters for example, devote their life to the art. Max Hammerton presents another of these exceptions of the natural selection when he asks: why is human intellectual ability so enormously variable? Galileo remarked that the difference between human beings in intelligence seemed greater than that between humans and other animals. (5)
A second alternative to the algorithm X is the randomness. But it is straightforward to realize that the randomness itself cannot be the source of our non algorithmic mathematical understanding because in that case it would be magical. However this element could affect something in the behaviour of the computers. The Penrose's answer is negative because the randomness can be simulated with no problem in the computational algorithms, although the sequence of steps of the algorithm would be predicted at the execution level, for an 'external' observer the output is unpredictable. (6)
Now Penrose considers two cases in which random behaviour takes place: the chaos and the quantum mechanism. The hope is to find in these phenomena something more that pure randomness. At first he considers the phenomenon of chaos; it happens when in a deterministic system there are so many parameters to be considered with such a deep accuracy that it is impossible to predict the behaviour of the whole system. The weather forecast is a well-known example of this situation: the description of the initial conditions would require so many parameters in such a precise way that a correct prediction is impossible in practice. Penrose does not exclude that this kind of structure could contain anything approximating genuinely non algorithmic behaviour, he says that for the moment the only concept that we can take from chaos is randomness.
About the quantum theory, Penrose explains that the behaviour of the molecules respects the

deterministic laws (respectively Schrodinger equation and Newton and Maxwell's laws) at both the quantum level of molecules and the classical level of everyday experience. But in a middle ground level between the quantum level and the classical one, it behaves in a random way. That leads to the EPR (Einstein - Podolky - Rosen) paradox: we imagine a quantum reaction that sends two identical particles, A and B, in two opposite directions. Then we measure A to determine if its spin is right or left handed, this is the random element because in quantum theory a particle does not have a definite spin until it is measured. After the measurement the wave function of A will 'collapse' to a random spin of the two possible. Surprisingly, wherever B is at that moment, it instantly undergoes a similar wave collapse with an opposite spin, this phenomenon is known as state vector collapse (from the vector of possible states after the measurement we collapse to only one state) or wave function reduction. The unknown aspect of this paradox is how the two particles manage to stay connected. (7)
Also in this theory Penrose does not find that non algorithmic ingredient that could be relevant to our thought processes. However he thinks that future developments of quantum theory will yield to a proper understanding of the physical process underlying quantum state reduction. In particular he argues that effects not currently covered by quantum theory should play a main role in the vector state collapse, and in his opinion these could be the gravitational effects (where 'gravitational' is to be intended in the general relativity theory). They should be incorporated in the quantum theory to arrive at the concept of gravitational induced objective state reduction. (8)
My personal opinion is quite similar, I believe that paradoxes like EPR suggest with clearness that the present-know physical laws are not enough to explain completely the quantum theory. I think that the quantity of debates about the quantum theory show that, although it is a corrected and proved theory, lots of things still have to be explained. The EPR one should not be considered only like an illustration of how quantum mechanics violates classical intuitions but a clue of a truth that we cannot explain completely yet. For this reason, furthermore for the non computational behaviour that our mind shows in the mathematical and artistic subject and for the absence of other possible algorithmic explanations of our consciousness, like the seven different alternatives examined by Penrose show, I refuse the second viewpoint among the four before presented. I think that there are still some physical law to be found. Like the Newton laws are not enough to explain phenomena which happen at a high speed, near the speed of light, or at the level of the atoms, in the same way I think that we need another genius physicist to find out a completely new theory, which will not reply, but will complete, the quantum theory. Maybe, after that, understanding the human consciousness will be easier. I said 'maybe' because it seems to me not so obvious the connection between quantum theory and consciousness. It could be that the main reason why there are so many debates about the relationship between quantum theory and consciousness, it is that the difficulty to accommodate the concept of state vector within our usual understanding of what physical reality is provides us an interface between the physical world and the non physical concept of consciousness. (9) But this interface is not so easy to be "implemented" indeed the present explanations of the connection between quantum theory and consciousness are revealed to be weak, for example about the Penrose's "gravitational induced objective state reduction" that I mentioned before it arises a new problem: this kind of reduction should be non deterministic but computable because it could be represented like a stochastic process and it would be useless to find the non algorithmic element. Penrose sketched some concepts about the non computational features of quantum gravity but they are not well definite yet. (8) Hameroff tried also to find this connection in the microtubules of the neurons. He argued that the superposition of tubule states over many neurons could collapse for an individual elementary act of consciousness. (10) But this idea was strongly criticized by Grush and Churchland who noticed lots of weakness in the theory, actually many of the Hameroff's arguments are just conjectures without solid evidences. For example Hameroff envisages consciousness as involving a unity across different brain regions. How might consciousness be transmitted from one tubule to its neighbour or furthermore among different neurons, with a different voltage of the

neuromembrane, is really unknown, neither it exists the evidence that it is done. He also supposes that in the microtubules there is only pure water but there is no known mechanism for keeping out common cytoplasm ions so these kinds of impurities are an obstacle to the quantum mechanical properties. (11) Actually the main difference between Penrose's arguments and mine opinion is that I am fewer optimists. According to me the way to explain in an acceptable way the consciousness is still long, even if one could argue that psychologists and biologists have not treated the issue seriously until very recently and big progress have been made in the last twenty years. (12) Like Penrose I think it is true that appropriate physical actions generate awareness because an explanation, as in the mysticism, which does not lead any proof and that requires a blind faith, is unacceptable. However the fact that these physical laws exist it does not mean that we will be able to find out them, because like a cat will never understand that the shape in the mirror is itself we might be not enough intelligent to understand everything of the hidden mechanism of consciousness. Some concepts of the scientific present-day knowledge are already so complicated that only some experts can really understand them, so why should not a limit to our knowledge be found? But now we can suppose for a while that a day finally we will find out these hidden laws and we will able to explain the mechanisms of our consciousness. Does it mean that we will be able to create an artificial mind aware of itself? I am really pessimistic about that possibility. I am sure that we will not be able to reproduce the artificial intelligence with the present concept of computers because they cannot reproduce the non algorithmic element of our mind. But I also think that theories like the Deutsch's one about quantum computers are at least 'futuristic'. He refers to the theory, accepted by Penrose, of many worlds interpretation of the quantum phenomena (7) according to which there is a different parallel world for every one of the possible outcomes of a quantum event. If this theory was true the number of task that would be performed at the same time would be unlimited because the quantum computer could distribute its operation to copies of itself in other universe. (13) Probably we will see realized this theory sooner than we could imagine, although not in the real word but in a science fiction movie!

References

1. Roger Penrose - "Setting the scene: the claim and the issue". Pag.1
2. Roger Penrose - "Setting the scene: the claim and the issue". Pag.4-8
3. John Casti - "Confronting Science's logical limits".
http://www.fortunecity.com/emachines/e11/86/loglimit.html
4. Roger Penrose - "Setting the scene: the claim and the issue". Pag.16-19
5. Max Hammerton - "Professor Penrose's mind". http://www.tribunes.com/tribune/art99/ham3.htm
6. Roger Penrose - "Setting the scene: the claim and the issue". Pag.24-28
7. Martin Gardner - "Theory of everything" -
http://www.newcriterion.com/archive/23/oct04/gardner.htm
8. Harald Atmanspacher - "Quantum approaches to Consciousness" -
http://plato.stanford.edu/entries/qt-consciousness/#4.5
9. Evan H. Walker - "Quantum theory of consciousness" -
http://users.erols.com/wcri/consciousness.html
10. Stuart R. Hameroff and Roger Penrose - "Conscious events as orchestrated spacetime selections" - Journal of Consciousness studies (1996)
11. Rick Grush and Patricia S. Churchland - "Gaps in Penrose's toilings" -
http://mind.ucsd.edu/papers/penrose/penrosehtml/penrose-text.html
12. Bernard J. Baars - "Can Physics Provide a theory of Consciousness?" -
http://psyche.cs.monash.edu.au/v2/psyche-2-08-baars.html
13. Peter Weibel - "Virtual worlds: The Emperor's new bodies" -
http://www.aec.at/en/archiv_files/19902/E1990b_009.pdf - pag.11-12