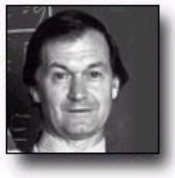
© *1997−2009, Millennium Mathematics Project, University of Cambridge.*
Permission is granted to print and copy this page on paper for non−commercial use. For other uses, including electronic redistribution, please contact us.

Jan 2002
Features

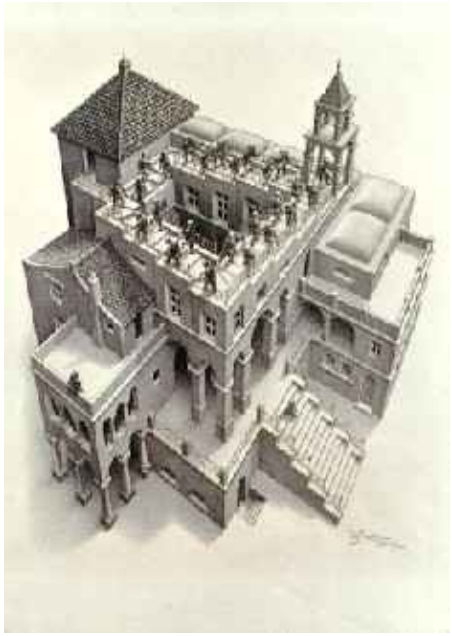


Roger Penrose: A Knight on the tiles

by The Plus Team

BOOKMARK

Sir Roger Penrose is one of the world's most widely known mathematicians. His popular books describe his insights and speculations about the workings of the human mind and the relationships between mathematics and physics. His interests range from astrophysics and quantum mechanics to mathematical puzzles and games. As a teenager, he invented the so−called "Penrose staircase", used by Escher in some of his famous optical illusion drawings, such as the one below. Helen Joyce from the Plus team talked to Sir Roger about his ideas.



M.C. Escher's "Ascending and Descending".
© 2002 Cordon Art − Baarn − Holland (www.mcescher.com).
All rights reserved. Used by permission.

Quantum Consciousness

Perhaps Penrose is best known to the wider public for his view that there is an essentially non−algorithmic element to human thought and human consciousness. Roughly speaking, an *algorithm* is a clearly defined set of steps for carrying out some procedure; once an algorithm for carrying out a procedure has been found, no more independent thought is required. All that is necessary is to follow the steps. Anything a computer can do must be algorithmic, since computers can only do what they are programmed to do, and programs are algorithmic in nature. This non−algorithmic element to human thought is, Penrose claims, due to quantum effects in the brain, which are the source of our feelings of self−awareness, our consciousness, and our capacity for leaps of inspiration. These ideas are set out very fully in Penrose's two bestselling works of popular science, "The Emperor's New Mind" and "Shadows of the Mind", the latter reviewed in this issue of *Plus*.

At the moment, the central project of mathematical physics is the search for a so−called "Grand Unified Theory" or "Theory of Everything". This theory would reconcile quantum mechanics, which operates at very small scales, and relativity, which operates at very large scales. These two beautiful theories both have much experimental evidence to support them, but unfortunately are incompatible.

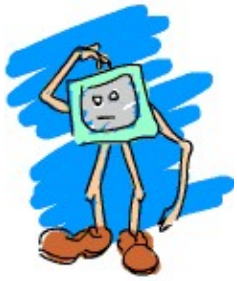
Roger Penrose giving a lecture at the conference [Photo and copyright Anna N. Zytkow]

Penrose is unusual in believing that quantum mechanics will have to change in order to fit into such a unified theory. He says that this is "what distinguishes my view from the others who believe that quantum mechanics are important in mental phenomena, although even that's a minority view. These other people don't, on the whole, think that you have to go that further step, and modify quantum mechanics. According to me, quantum mechanics, as we understand it today, has certain limitations and will be shown not to be correct. The level at which this will happen is around the limits of what is measurable with present−day technology. This is interesting, quite apart from its implications for consciousness and AI.

"The more detailed ideas I have about consciousness do depend on specifics − I must be right, not only in that quantum mechanics has its limits, but also in the detailed way in which it must be changed. So if the experiments I'm trying to get colleagues to perform show that I'm wrong then my model of consciousness will also be disproved."

Penrose's ideas on consciousness are, to say the least, controversial in the AI community. After all, what he is saying is that the goal of an intelligent, let alone conscious, computer is unattainable right from the start, a view understandably unpopular with those who are dedicating their lives to creating such entities.

However, Penrose's opinions are given some force by the slow progress of work on AI. It cannot be denied that early proponents of AI grotesquely underestimated the difficulty of their task. Arthur C Clarke placed HAL, his thinking, feeling, ultimately paranoid computer, in 2001, during which year real AI researchers were teaching small robots to find their way around rooms! As Penrose says, this is "no indication of any sort of intelligence, let alone consciousness".



I am, therefore I think?

Penrose also disagrees with the majority view regarding how to decide whether a computer is conscious. For over fifty years, since Alan Turing wrote his hugely influential article proposing the so−called "Turing test" as

a way of deciding whether a machine is conscious, most workers in the field of AI have accepted, in effect, that if a machine can persuade a human observer that it is conscious, then it must be.

To be more accurate, Turing proposed that an interrogator in a separate room should interact with a computer and a person, via a machine terminal. If, after some unspecified time, the interrogator could not tell the difference between the computer and the person, the computer should be regarded as "intelligent".

Not only does Penrose feel that the Turing test bypasses the essential element − consciousness − but he also believes that a non−conscious machine couldn't "fake it". In his opinion, a really good impersonation of consciousness by a machine would not be possible unless the machine actually were conscious. He says that "within a limited framework, one might get a performance that would fool somebody for a while, but ultimately the machine would give itself away. Of course there's a question just how good your simulation might be, and, so far anyway, the people working in this field are nowhere close."

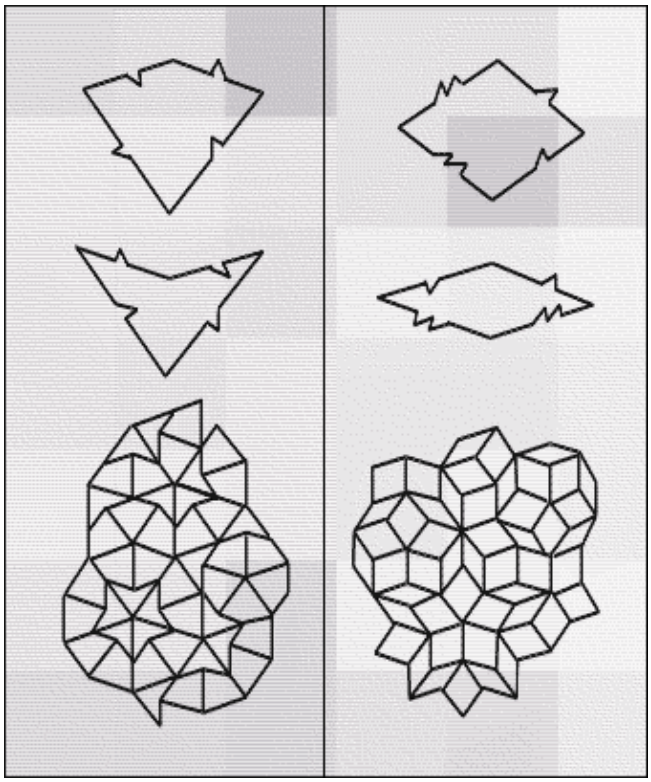
Trouble with Tiling

A problem from recreational mathematics to which Penrose has made a significant contribution is the *tiling problem.* His work on this problem is described in some detail in Quasicrystals and Kleenex, an article from Issue 16 of *Plus*. The tiling problem is this: given a collection of polygonal shapes, is it possible to cover the whole plane using just these shapes, with no overlaps? Such an arrangement of shapes is called a *tiling.* Tilings are said to be *periodic* if they are exactly repetitive in two different directions (a direction and its opposite are not counted as different!).

At first sight, the problem of whether one can find a tiling of the plane by shapes which will *only* tile non−periodically, seems lighthearted. In fact, although it is "fun" mathematics, it has a philosophically deep aspect − it is part of the area of mathematics known as *non−recursive.*

A class of mathematical problems is called *recursive* if there is an algorithm for finding the answer in each individual case. It is called *non−recursive* if there is is a yes−or−no answer to each individual problem, but there is no algorithm for deciding whether the answer is yes or no in each individual case.

It has been known for about forty years that there is no algorithmic way of deciding whether a given collection of polygonal shapes will tile the plane, that is, the tiling problem is non−recursive.



Two Pairs of Penrose Tiles

The way in which this was originally proved was interesting. First it was shown that if, whenever a collection of polygonal shapes, or "tiles" could tile the plane in some way, the same collection could tile the plane periodically, then the tiling problem had to be recursive. The next step was that the tiling problem was proved not to be recursive. This meant that there had to be a collection of polygonal shapes that could tile the plane, *but only non−periodically*. The first such collection of tiles that was found was was absolutely enormous − it contained 20,426. Various mathematicians found smaller collections, culminating in Penrose's discovery of a pair of tiles that tile the plane only in a non−periodic way. In fact, there are more than one such pair − two are shown below.

It's interesting to think about the problem of inventing a tiling in the light of Penrose's thoughts on AI, consciousness and inspiration, because the problem seems exactly like the sort of thing a computer couldn't do. Penrose says that people ask him where he got the inspiration, whether it came to him suddenly.

"It wasn't like that, but, you see, it's never like that. What it's like is that one has been thinking about a problem for a long time and getting very familiar with it. After a while, you may be worrying about some particular aspect of the problem, and then it's quite possible, even when you're thinking about something else, that an idea may come to you, and you realise that things fit together in some way, with the accompanying realisation that this must be right. Of course sometimes one may be mistaken in such things!

"You see, I was already aware of certain nonperiodic patterns of pentagons and other shapes which are just nice to look at. What you might call the inspiration was to realise that, by modifying the shapes in different sorts of ways, by putting knobs on like jigsaw pieces, you could force the pattern by local matching rules. As a pattern, I was aware of it before, but the thing that required a little bit more, something for me to guess or something to come from somewhere else, was that realisation."

As Penrose says, one of the interesting features of the tiling problem is that it is noncomputable. "It is possible to explain noncomputability in relatively simple terms, but people almost always get the wrong end of the stick! They say, 'show me a problem you can't solve', but it's never like that! It's a *class* of problems which have no systematic solutions. With the tiling problem, I sometimes say in talks, 'well it depends on the existence of nonperiodic tilings', and give an example of such a tiling. And then at the end of the talk people come up to me and say, 'how do you know that that tiles when the problem is noncomputable?' but that's not the point. The point is that in a particular case you may have a way of seeing the solution, but there is no systematic procedure that could be put on a machine, which requires no more thinking."

An analogy that springs to mind is the fact that there are uncountable sets, such as the real numbers. A set is said to be *countable* if its elements can in principle be listed. An example is the natural numbers, 0,1,2,3,... We are sure that the list that we get by adding one to the last element to get the next element will give us all the natural numbers.

In fact, the rational numbers are also countable (rational numbers are fractions). But the real numbers − all the numbers on the number line − are not. When people first learn this fact, they are prone to think that individual real numbers are uncountable, and to ask "what is the number you can't count?", or, "What's wrong with root 2?"! But that's not the point. We could add any individual real number to the set of rational numbers and still have a countable collection, just as we may be able to prove individual problems from a non−algorithmic class of problems.

Discovery or Invention?

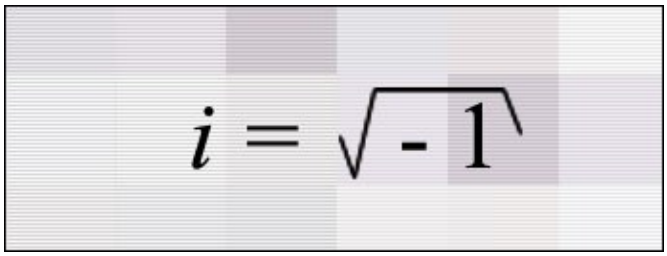
Penrose shows himself an unabashed realist, by proclaiming that acting conscious is not the same as being conscious. His realism is nowhere more evident than in his thinking on the source of mathematical inspiration. Although he takes a nuanced view on the hoary old question of whether maths is "out there" or "in here", discovery or invention, he is sure that it is not entirely a construct of the human mind.

"I've always thought that the distinction between invention and discovery is not that clear. Just the other day some people were talking about this sort of thing and Edison was mentioned, about him trying all these different things to make his first light bulb. They were saying that this was clearly invention.

"But it seems to me that we can't say that, because if the right substances weren't out there for him to use he couldn't have made it. So it clearly had elements of discovery as well as elements of invention.

"In mathematics it's often not too dissimilar to that. In a sense the big results are out there. I think this is the way to think about it, but in your access to these, if you need to prove some sort of theorem which is a stepping stone to some result, there may be all sorts of different ways you could go. These will be very much of the invention type. So I think that invention plays a big role in mathematical research, but there is nevertheless an element of discovery.

"This discovery element is most clearly visible in those areas of mathematics where you get out so much more than you put in. The biggest example of that that I know of is *i*, the square root of minus one.



"You introduce this number in order to solve just one equation, and then suddenly you find that you can solve all these other equations, which you had no conception of at the time.

"Nothing like that happens when Pythagoras' Theorem leads you to introduce the square root of 2. You have to go beyond the rational numbers to find the square root of 2, but if you just add the square root of 2 to the rational numbers, you get nowhere, you don't even get the square root of 3!

"You have to have the real numbers there first, but then *i* gives you a whole new universe. A window has opened into a completely new world."

A view into the new world of complex numbers

Plus is part of the family of activities in the Millennium Mathematics Project, which also includes the NRICH and MOTIVATE sites.