

The Penrose Institute™

*The Roger Penrose Institute for the Study of Creativity, Consciousness and Cosmology*

# Contents

[Introduction to the Institute](#)

[Sir Roger Penrose](#)

[The Three Areas of Research](#)

[I. Human and Artificial Intelligence](#)

[II. Quantum Biology Laboratory](#)

[III. New Physics & The Origin of Our Universe](#)

[Mathematical Play](#)

[Aging and Cognitive Longevity](#)

[The Institute Plan](#)

[Collaboration Methodology](#)

[Structure and Funding of the Institute](#)

[Visioning 5 years in the future - what did we discover?](#)

[Why San Diego](#)

[FAQ](#)

[Appendices](#)

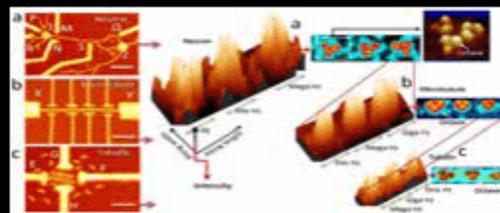
# Introduction to the Institute



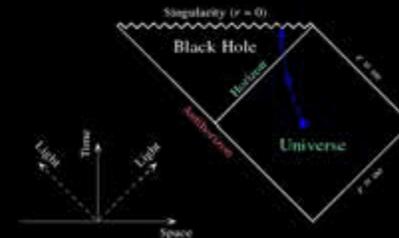
We are entering a new scientific era where the unification of quantum mechanics and general relativity will create new physics and new devices. 30 years ago Roger Penrose wrote a seminal book – *The Emperors New Mind*. Unification, he said, would require modification to quantum mechanics and that modification would explain human consciousness. Scientists met his proposal with scepticism. Quantum phenomena were only known in laboratories near absolute zero. But, recent breakthroughs show quantum phenomena are important in biology at room temperature. Photosynthesis gets its efficiency from coherent transport of energy, and some birds have a quantum compass located in their eyes. We believe studying unification will throw light on the fundamental laws of physics and generate technology in the fields of AI, human disease and aging, along with developing new types of incredibly sensitive sensors to probe the universe and our planet.



Human Intelligence



Quantum Biology

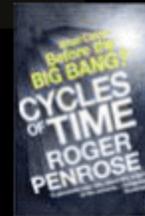
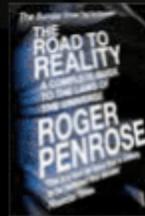
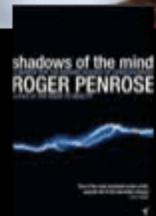


New Physics

# Sir Roger Penrose OM FRS



Sir Roger Penrose, Emeritus Professor at the Mathematical Institute of the University of Oxford, Emeritus Fellow at Wadham College, and winner of the Wolf Prize in Physics has made profound contributions across a broad range of scientific disciplines. His work encompasses geometry, black hole singularities, the unification of quantum mechanics and gravity, the structure of space-time, and the origin of our Universe. His geometric creations inspired the works of Escher, and the Penrose steps have been featured in several movies. His tilings adorn many public buildings, including the Oxford Mathematics Institute (pictured), and will soon decorate the San Francisco BART system. The five fold symmetry, initially thought impossible or a mathematical curiosity, has now been found in nature. In 1989 Penrose wrote *The Emperor's New Mind* which challenged the premise that consciousness is computation and proposes we need new physics to understand it.



# Areas of Research

## I. Intelligence (Human and Artificial)

Human intelligence appears to be very different from today's artificial intelligence. But, artificial intelligence is getting stronger by the day. Is there something human beings can do that an artificial intelligence cannot do? This is both a fundamental question in mathematics and a question with enormous social implications as we decide what people should learn in the next few decades, and how we should best work with artificial intelligence in the future.

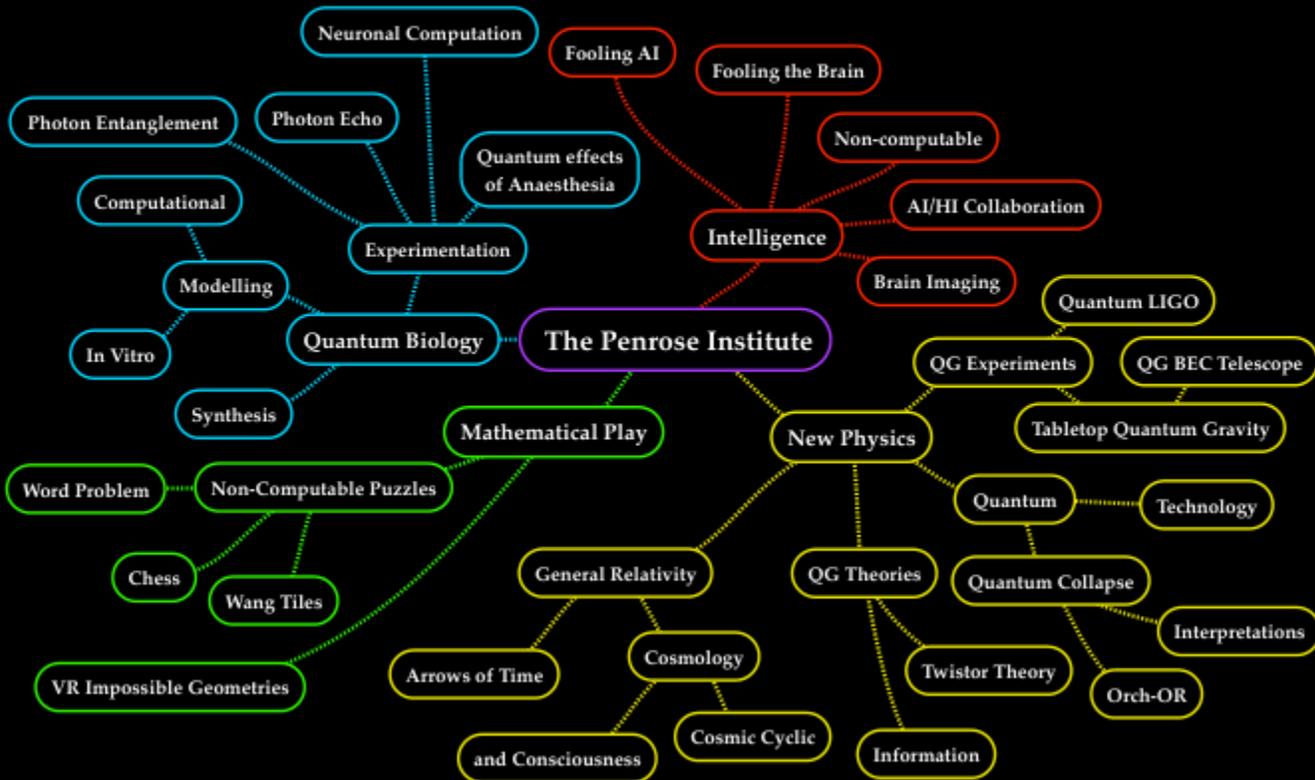
## II. Quantum Biology Laboratory

It used to be thought quantum mechanics played no part in biology and was confined to experiments near absolute zero in physics laboratories. A revolution is taking place. We now understand quantum mechanics is important in modelling chemical processes in the body, such as, protein folding in the presence of water. More excitingly, we have seen the first indications of exotic quantum effects. Photosynthesis gains its efficiency from quantum coupling and some birds appear to use a quantum compasses located in their eyes. How important are quantum effects in the body and, do we need new physics to explain the working of the brain?

## III. New Physics and the Origins of our Universe

We will investigate the interplay of quantum physics, general relativity and information theory. Roger Penrose's work on Quantum gravity (twistor theory) and conformal cyclic cosmology along with non-causal computing already open new theoretical approaches. We will explore the possibility of incorporating recent developments in quantum information, quantum biology and quantum thermodynamics to build a bridge between the physics of the brain and consciousness. Theories will be tested in quantum systems such as Bose-Einstein Condensates (BECs). New physics might not only help us understand human consciousness, but also provide a deeper understanding of the origin of the Universe itself.

# Institute Concept Map



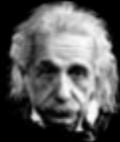
# Scientific Philosophy

The institute will follow the scientific philosophy of Roger Penrose across a broad range of interrelated scientific domains. This philosophy is born of a passion to uncover paradoxes in our current understanding of physics and propose new ideas that can be experimentally tested. Such tests may, of course, result in the disproof of the ideas, even his own! Our principles:

- Uncover paradoxes and failings in our current physics theorems.
- Consider both the physical world and the minds which observe it.
- Be creative, novel and challenge the status quo.
- Follow a rigorous and detailed scientific method with falsifiable experiments.
- Be unfashionable, brave and controversial without being discourteous.
- If a principle needs to be changed, explain why and change it!

# I.

## Human & Artificial Intelligence



*"I never made one of my discoveries  
through the process of rational  
thinking" – Albert Einstein*

## The Intelligence Question

At the heart of the artificial intelligence debate is a fundamental disagreement between two theories of the human mind. Proponents of strong AI argue the human brain is a computer. As computers become ever more powerful, following Moore's law, they will eventually eclipse human intelligence, perhaps even this century. On the other hand, many scientists and philosophers believe the human brain is not a computer. Explaining why is hard. The human mind appears to be the only structure that uses conscious action to create new things. Such creativity might surpass the capability of a Turing Machine and this is the belief of Roger Penrose. However, AI technologists are making progress in building computers that appear to display creative intelligence. This is a live debate.



**Our Goal:** To examine creative thinking through practical experimentation and theoretical work so that we might understand how the brain achieves its power, improve that power and better inform our ability to work with AIs in the future.

## Impact of AI on Society

Over the next few decades AI will profoundly change the nature of human work. What should humans do with our apparently unique gift of creative intelligence. Is this gift truly unique?



© The Penrose Institute 2016-17

10

# Finding Creativity in the Brain

We will identify people with exceptional abilities at solving non-computable puzzles, image and record their fMRI, MEG and EEG activity during the process. We hope to identify brain regions and EEG frequencies and patterns associated with non-computability. These patterns will then be used as feedback to enhance and train non-computable creative processes.



As an example, we might have, for our initial list:

EAT = AT  
ATE = A  
LATER = LOW  
FAX = FELLOW  
CARP = ME.

From these we can derive, for example,

LAP = SEAP



Recruit people who can solve non-computable puzzles in a similar fashion to the way code breakers were recruited in World War II.

Offer them a combination of computable and non-computable puzzles to work on.

Image the brains using MEG, EEG, fMRI and others technologies to see which areas of the brain are involved.

# Insights, Outcomes, Spin-outs

## Theoretical

- Locate the boundaries between computable intelligence and creative intelligence.
- Research how human brains might be creative when this would normally be a non-computable operation – computational ‘tricks’ or non-computable operation.
- Demonstrate how non-computable / non-causal models of computation might go beyond current Turing architectures.
- Research fundamental questions of mathematics such as determinism.
- Determine whether non-Turing computable proofs have Gödel incompleteness.
- Develop new objective forms of Turing Tests.

## Experimental

- Locate the ‘seat’ (or process or pattern) of creativity within the brain.
- Create better interfaces between human intelligence (HI) and artificial intelligence.
- Develop extended computational models such as non-causal computing.
- Investigate the inheritance of creativity

## Spinout / Partnerships

- New forms of efficient computational systems.
- Artificial/human intelligence collaboration systems.
- Training and improving creativity

# II.

## Quantum Biology Laboratory

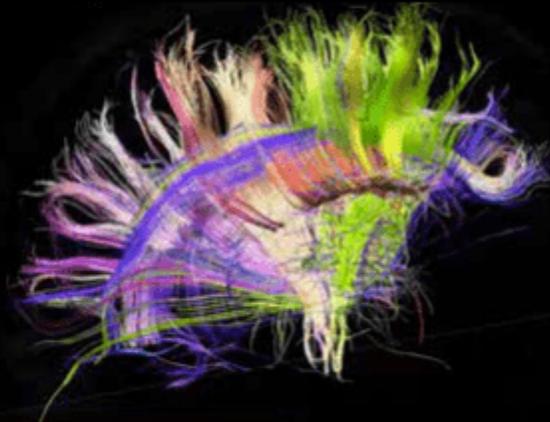


*“And you are made of a hundred trillion cells. We are, each of us, a multitude.” – Carl Sagan, Cosmos*

# The Quantum Neuron?..

Current studies focus on the connectome – the map of connections between the neurons. Sophisticated maps are being built by governments, including the US BRAIN Initiative and the EU Human Brain Project (HBP), along with private institutes such as The Allen Brain Institute. The EU project aims to simulate the brain by running a model of the connectome on super-computers. However, problems arise as a pure connectome model does not appear to capture sufficient information to model the seemingly simple neural networks such as the the *C. elegans* nematode worm. Recent discoveries suggest quantum effects might be significant in brain tissue.

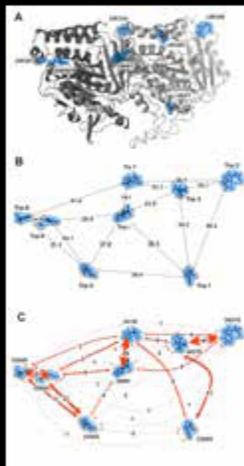
**Our goal:** To probe and model the sub-neuronal structure of the human brain and determine whether, and in what manner, small scale and quantum effects are significant to its operation. Understanding how quantum effects might be a factor in neurons would add beneficially to human knowledge and help in all manner of practical problems, from improving human cognition to curing brain diseases such as Alzheimer's and lead to new forms of computing machine.



Van Welden, ..., Martinos Center and Dept. of Radiology, Massachusetts General Hospital and Harvard University Medical School

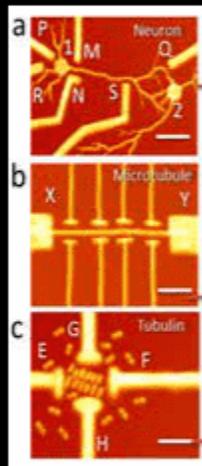


# Quantum Brain Tissue Experiments



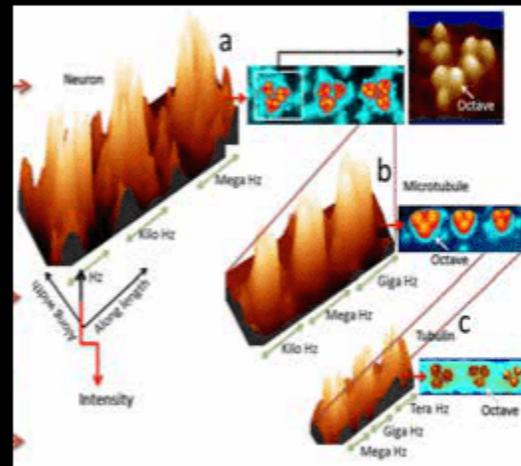
Quantum Modelling

Travis Craddock et. al. Nova

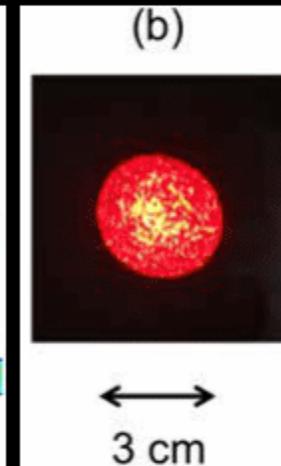


Quantum Imaging

Anirban Bandyopadhyay group Sahu et al, 2013a; 2013b; 2014



Quantum Resonance & Transmission



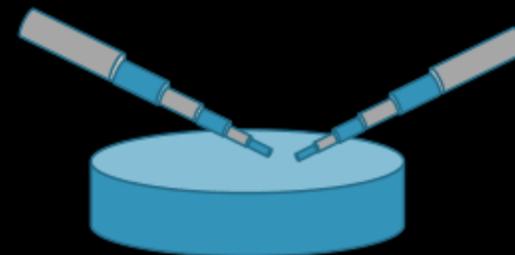
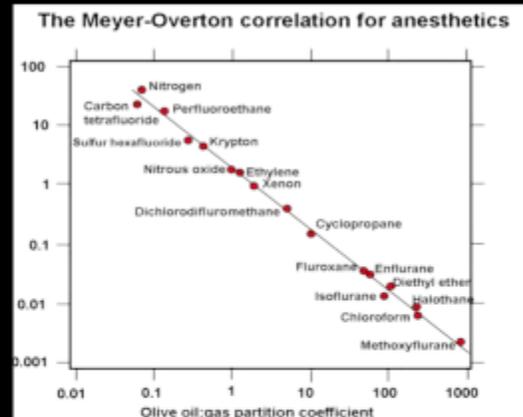
Quantum Entanglement

Photon Entanglement Through Brain Tissue  
Lingyan Shi<sup>1,2,3</sup>, Enrique J. Galvez<sup>4</sup> & Robert R. Alfano<sup>1</sup>

# Un-Consciousness in Vitro

In the last year it has become possible to test living tissue at the nano scale so that we can uncover quantum effects. It is our intention to apply these technologies to attempt to understand consciousness.

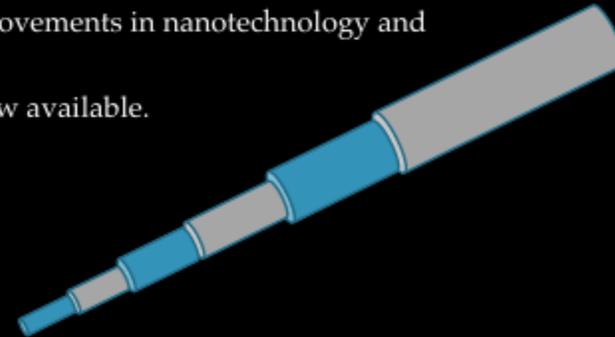
- Our main probes for consciousness are anesthetics. We can anesthetize a human subject and they will appear to go unconscious and report that they did so once reawakened. There are two main gaps in our understanding. We do not understand how anesthetics work and we don't have an objective measure for unconsciousness, let alone consciousness!
- Anesthetic effectiveness correlates with solubility of the compound or gas in olive oil, suggesting polar effects are significant. It is unknown where the anesthetics act to give their effect.
- Our experimentation will involve anesthetizing brain tissue and looking for key quantum effects such as bulk conductivity, resonance and preservation of quantum information.



# Nano-Brain-Technology

Our approach is made possible due to recent improvements in nanotechnology and modelling which we intent to further enhance.

- Tri-axial 0.5nm probe tips (patch clamps) are now available.
  - Relatively unbreakable
  - Repeatable probing
  - True ion channel resolution
  - Broad frequency range
- High speed optical cameras
  - High speed images can be taken across a broad visible spectrum.
- Neuron Culturing for Human, Neanderthal and Animal models
  - A range of different neurons can be generated from stem cells
  - DNA segments can be spliced into neurons to vary their genetic makeup
  - Culture can be kept alive for several months allowing training and characterization
- High fidelity molecular modelling
  - Proteins can now be computationally modelled with an accurate depiction of water
  - Quantum computing is beginning to become available to perform the modelling



## Locating Neuronal 'Creativity'

We are able to culture a range of neurons with features from different species, including Neanderthals – man's closest relative. In their several hundred thousand year existence they did not develop language or tool use. It is suggested this failure occurred because they devoted larger portions of their brain to motor control. However, it might be better explained through difference in the fundamental working of their neurons. This might allow us to locate the fundamental neuronal structures which explain homosapien creativity and language.



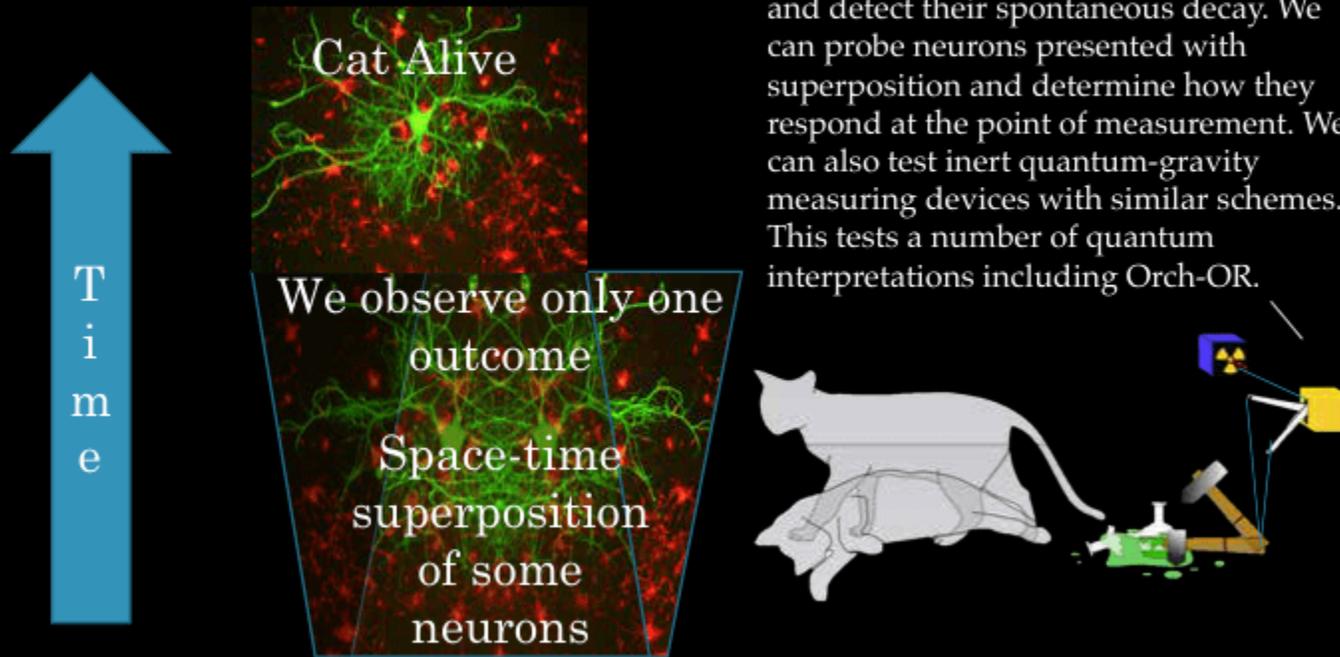
**Human**



**Neanderthal**

# Quantum Measurement (Orch-OR)

At Some Point our Neurons Must Perceive the Cat as Dead or Alive



We can build highly sensitive quantum systems to keep photons in superposition and detect their spontaneous decay. We can probe neurons presented with superposition and determine how they respond at the point of measurement. We can also test inert quantum-gravity measuring devices with similar schemes. This tests a number of quantum interpretations including Orch-OR.

# Solving the Halting Problem with Biological Neurons

## Computation with Neurons

- Human neurons are the most efficient and effective computational elements known. They are the fundamental basis of consciousness and creativity. Can we build a computer out of them? Even better, can we build a system based on human neurons to carry out non-computable operations?
- The fundamental non-computable operation is Turing's "Halting Problem" (Entscheidungsproblem). In problems of non-computable complexity, a machine might never halt: this is the unexplored capacity of the human brain.
- Can we build a network of living human neurons that can be programmed? Can we program that network to attempt a non-computable problem? What will that network do?
- Our contention is that we can build the network, program it and introduce a non-computable problem. Our greater contention is that programming the Halting Problem will NOT result in an infinite loop. The neuronal network will detect the infinity and stop action.
- The programmable neural network will have a proto-consciousness and the elements necessary for non-computable operations.
- These networks will be the basis for synthetic neuronal structures of a type never seen before and a model for study of drugs, stimulation control and other research and interventions for medicine.

# Aging and Cognitive Longevity

Research into neurons has several applications to aging.

Neurons are one of the few cells in the body that do not age. For the most part the cells we are born with are cells we die with. Unlike many body cells which are simply recycled these cells have developed robust methods for repairing themselves. Understanding these mechanisms may yield valuable insight into the aging process.

Unfortunately, the repair mechanisms which can keep cells working into our late nineties are subject to degradation and disease. San Diego has two leading Dementia Research centers for Alzheimer's. It is now thought that a major factor in these diseases is the break down in the cytoskeletal structures inside the neurons, the self same structures that Roger Penrose and Stuart Hameroff believe provide us with creativity, understanding and ultimately consciousness.

On top of the sub-cellular repair mechanisms, our brains seem to repair themselves, or at least, protect their function if we exercise them. Brain training software appears to help brains function better and we would like to understand whether exercising the brain with non-computable / creative tasks is more beneficial than rote type learning tasks.

# Insights, Outcomes, Spinout

## Theoretical

- Molecular modelling of neurons and the study of how anesthetic gases interrupt their operation.
- Construction of a programmable computer synthesized from living neurons will provide insight into the operation of the human brain.

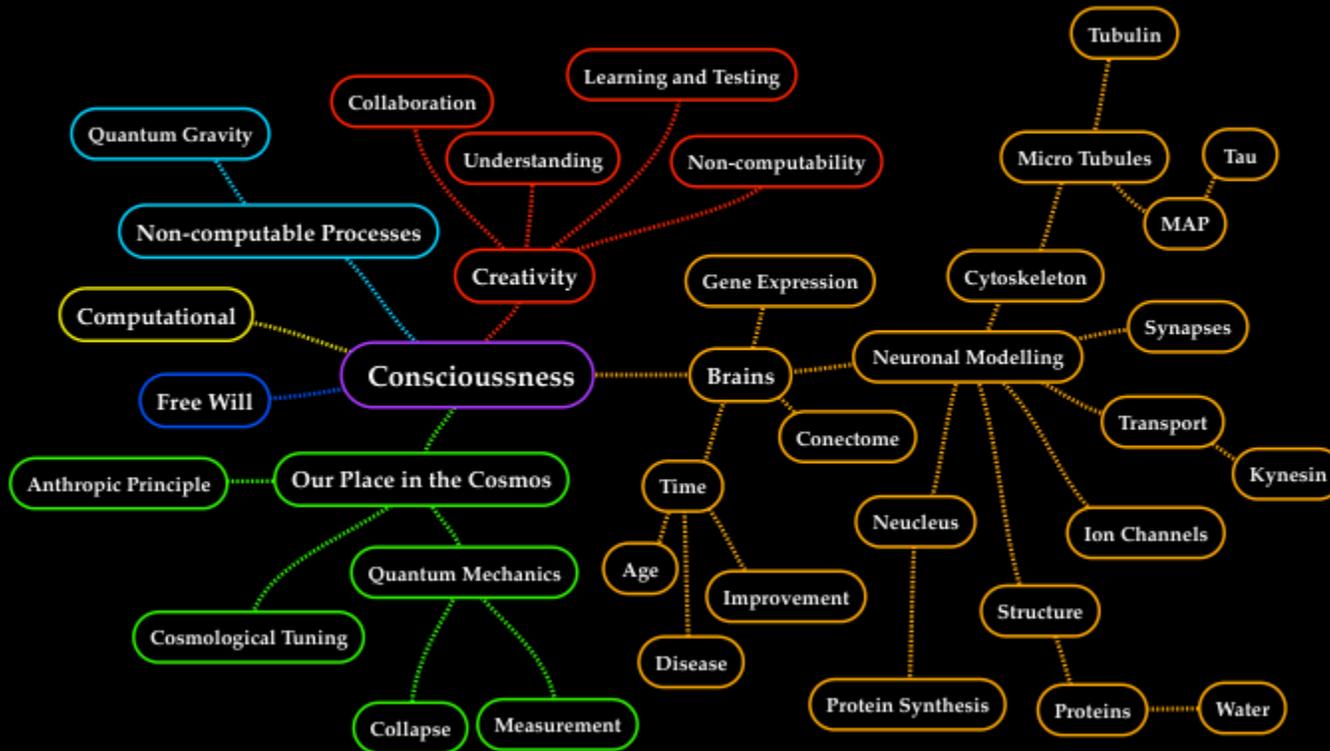
## Experimental

- Develop safer anesthetic gases through testing and certification in the models
- Find the key physical elements of a new quantum gravity computing technology.
- Use living networks of human neurons to serve as a means of screening medications for toxicity and their effectiveness in modifying neuronal actions.
- Study pathologies of neurons: accumulation of proteins in Alzheimer's Disease, Parkinson's Disease or damage from low-oxygen, such as hypoxia in strokes.
- Investigate non-invasive brain stimulation methods via the eye or transcranially.

## Spinout / Partnership

- Better enhance creativity and defend against dementia.
- Pharmaceutical screening capability with the neuronal network model.
- Contributions to computational science, including management of non-computability and new computing architectures, will be seen with the living neuron computers.

# Quantum Consciousness



# III

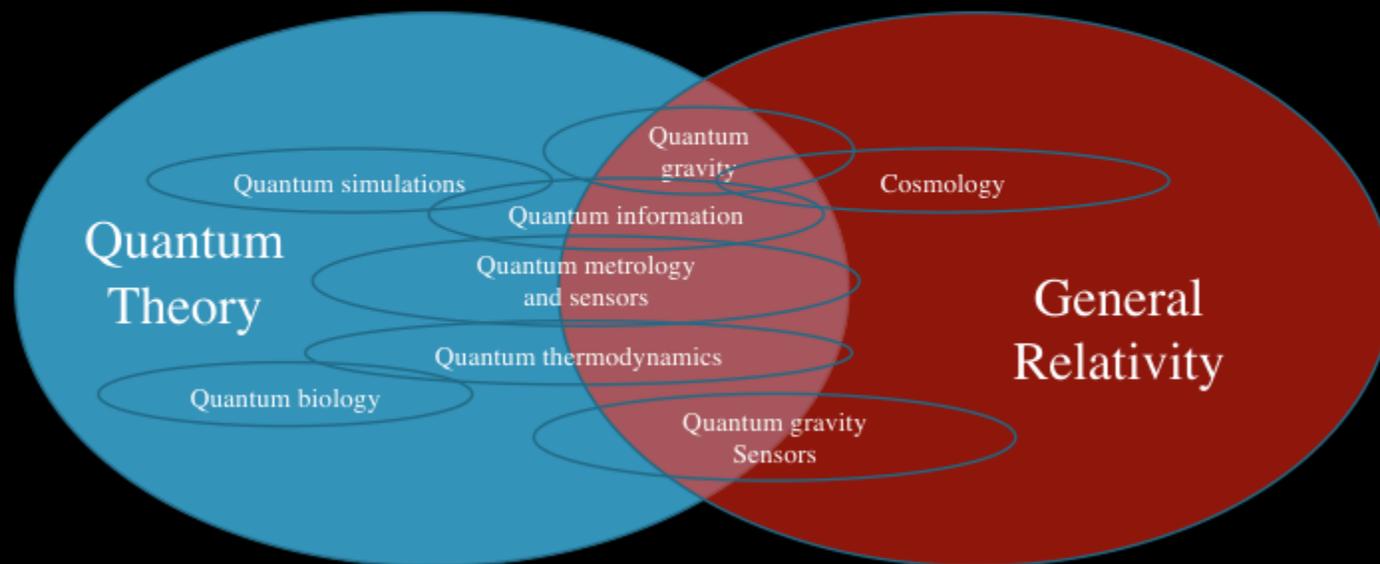
## New Physics

and The Origin of Our Universe



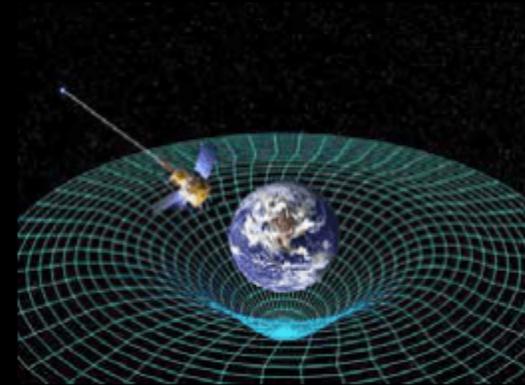
*"I think nature's imagination is so much greater than man's, she's never going to let us relax" – Richard Feynman*

# Research at the interface of quantum theory and general relativity



## Physics at the overlap of quantum information and relativity

Research at the overlap of quantum physics and information theory has given rise to the field of quantum information and its applications included quantum computation, cryptography and communication. This has led to a deeper understanding of the role of information in the physical world and to a new technological era. However, the world is not only quantum but also relativistic. Only recently, the use of relativity in quantum technologies has been considered and it has been shown that quantum systems can be used to measure gravitational effects. It is very likely that this body of work will play a role in our understanding of the physics of the brain. Penrose's work has already hinted at the key role that gravity might play in quantum physics and therefore, in our understanding of nature from neurons to cosmology.



## Quantum Gravity Sensors

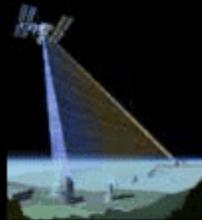


- Quantum systems, such as BECs, can be used to test the effects of gravity on the collapse of the wavefunction.
- BECs can also be used to measure gravitational effects such as gravitational waves, space-time parameters and perhaps dark energy and dark matter.
- Applications: quantum sensors, clocks and gravimeters compatible with notions in general relativity.

# Desktop Gravitational Wave Detector

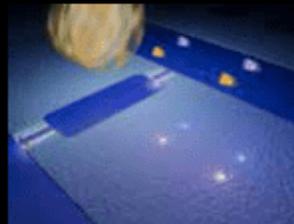
We are rapidly moving from a situation where Quantum Gravity effects are the realm of big-science, CERN, LIGO et. al. to a state where sensitive, innovative laboratory experiments can discover interactions and throw light on different QG theories.

- Quantum experiments are reaching relativistic regimes in space-based settings, fast moving boundary conditions and their interactions with quantum fields and in extremely precise measurement technologies and clocks. However, our understanding of physics at these regimes is very limited.
- Cutting edge experiments promise to deepen our understanding of quantum physics at the regimes where relativity can no longer be neglected also serving as testing grounds for Penrose's theories.



Quantum state transmission across thousands of km using satellites. At this regimes relativity kicks in.

(Zeilinger, Villoresi, Marquardt, et. al.)



Quantum fields interacting with boundaries moving at a third of the speed of light. (Delsing, Wilson)

Time dilation measured by quantum clocks separated by a few cm. Such clocks keep time to within 1 second in about 3.7 billion years. (Wineland)



# Quantum-Gravitational Telescope

**BECs, Bose-Einstein Condensate sensors, the size of a silicon chip and highly sensitive to gravitational geometry and waves.**

## Look Down with a BEC sensor

- Image the Earth's gravitational field.
- Use many to resolve feature
- Detect gravitational anomalies
  - Oil reserves
  - Map Earth's mantle
  - Fault lines
  - Fault movement
  - Ocean floor movement
  - Volcanic uplift



## Look Up with a BEC sensor

- Image the sky with gravity
- Make a gravity map of space
- See before the Universe was optically transparent 380,000 years old.
- See gravitational events
  - Neutron star Mergers
  - Black hole mergers

# Deep Space Gravitational Image

What might we see in the gravitational 'spectrum'

Cosmic  
Gravitational  
Background,  
'Gravitational  
isobars'

Merging Neutron  
Stars

Merging Black  
holes

Gravitational  
Voids

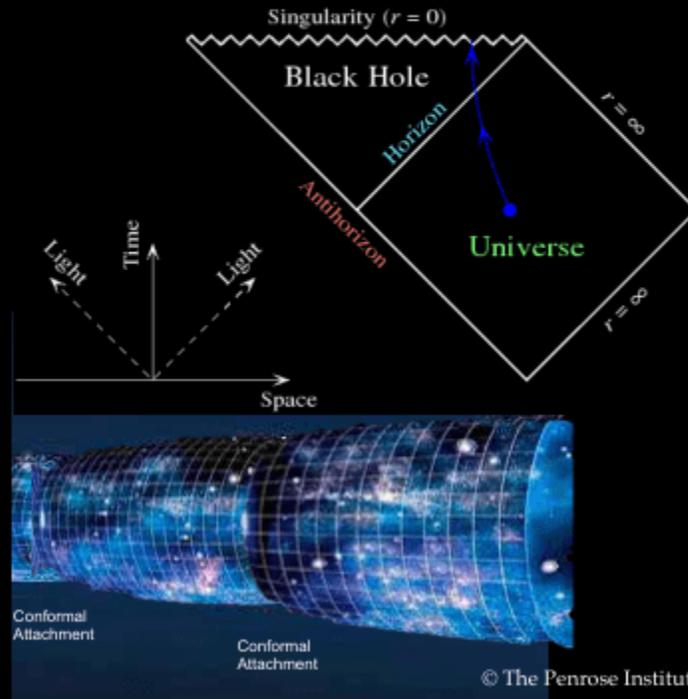
Gravitational  
Lenses

Dark Matter  
Anomalies

Artist impression based on Hubble eXtreme Deep Field, [NASA](#), [ESA](#), G. Illingworth, D. Magee, and P. Oesch (University of California, Santa Cruz), R. Bouwens (Leiden University), and the HUDF09 Team

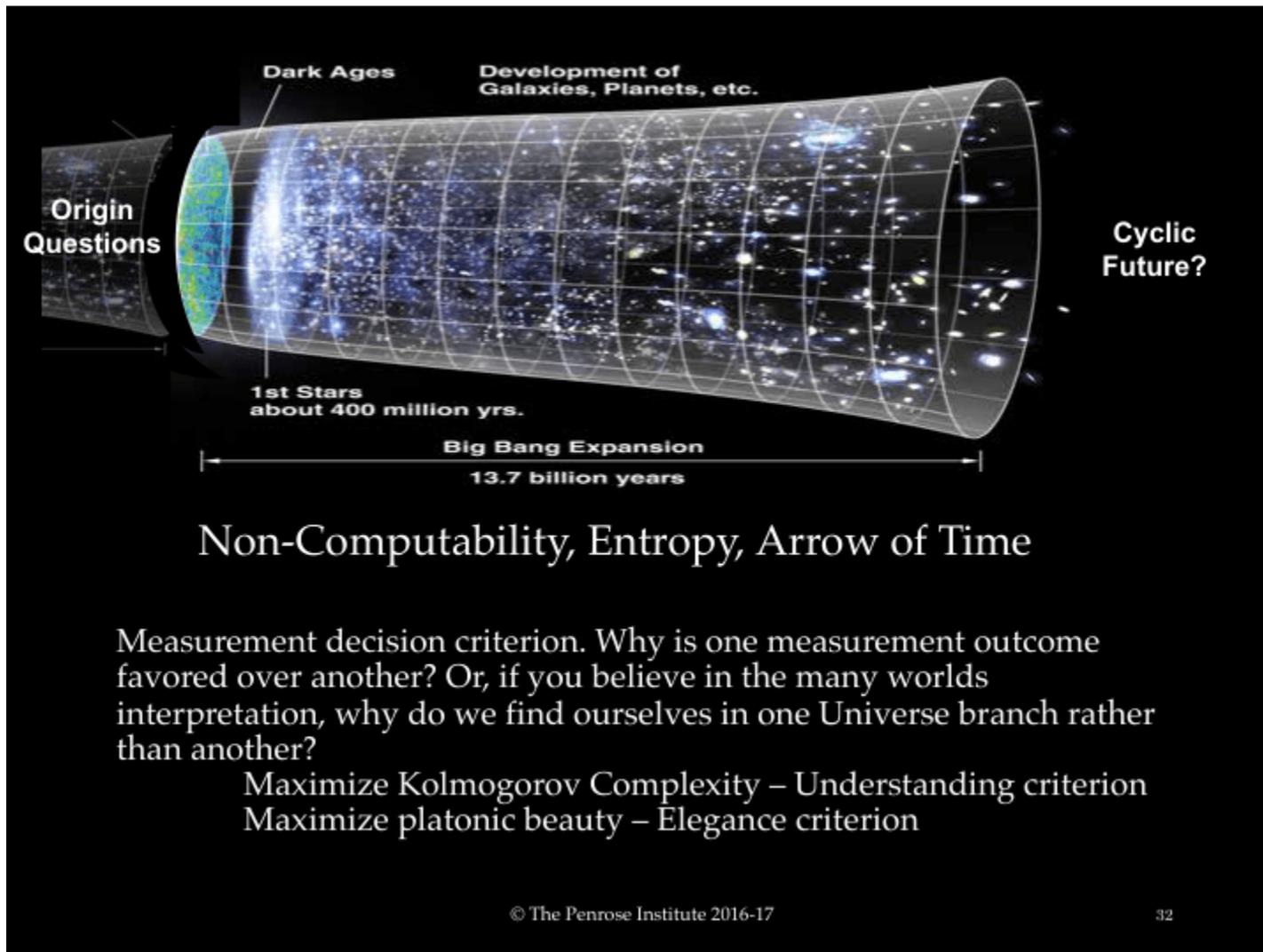
# Cosmology

In 2000, *Nature* asked 10 prominent physicists for their view of a 'Theory of Everything'. Among them, only Roger Penrose included a role for consciousness in the universe.



## Problems in Scope 2017

- Unification of QG
- Twistor Theory
- Conformal Cyclic Cosmology
- Dark Matter and Energy
- Background Independent Quantum Mechanics
- Non-causal Physics & Computing
- Computational Universe
- Conservation of Information
- Testing OR collapse by  $E=h/t$
- Arrows of time
- Low Entropy Origin



# Insights, Outcomes, Spinout

## Theoretical

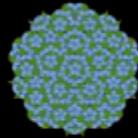
- Make progress on unification quantum gravity and general relativity.
- Develop insight into the implications of non-computability on the evolution of our Universe and the arrow of time.
- Understand how the brain might utilize quantum gravity effects.

## Experimental

- Build a quantum gravity interferometer using a Bose-Einstein condensate.
- Investigate patterns in the cosmic microwave background radiation.

## Spinout / Partnerships

- If it can be demonstrated that quantum gravity is implicated in consciousness then new forms of quantum gravity computers should exist and it would be possible to build one.
- Bose-Einstein condensate sensors would be the most sensitive sensor known to man. The last time we developed a new sensor of this sort – the superconducting quantum interferometer – it has application in a broad range of fields.



## Mathematical Play

*"We have a closed circle of consistency here: the laws of physics produce complex systems, and these complex systems lead to consciousness, which then produces mathematics, which can then encode in a succinct and inspiring way the very underlying laws of physics that gave rise to it." — Roger Penrose (The Road to Reality: A Complete Guide to the Laws of the Universe)*



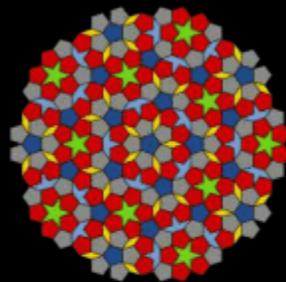
Penrose Tiles at the Mathematics Institute, Oxford



# Inspirational Mathematics



Illusions



New Forms



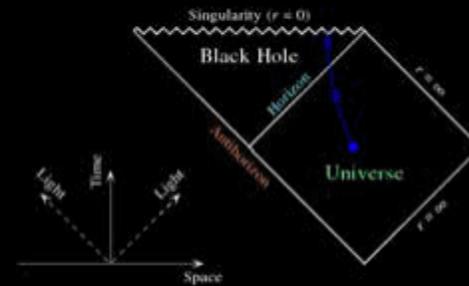
Impossible Shapes



AI/HI Puzzles



Art Inspiration



Infographics

# A New Turing Crossword Puzzle

If you can solve this puzzle in under six minutes, science needs you!

**WIN A TRIP TO LONDON**  
DO YOU HAVE WHAT IT TAKES TO CRACK THE CODE?  
COMPLETE ALAN TURING'S CROSSWORD PUZZLE  
(ORIGINALLY PUBLISHED IN THE *DAILY TELEGRAPH*)  
AND BE ENTERED TO WIN A TRIP FOR TWO TO LONDON

**INSPIRED BY THE IMITATION GAME  
STARTS TONIGHT!**

2			6	7	5		
							9 6
6	7			1	3		
5		7	3	2			
7							2
			1	8	9		7
		3	5			6	4
8	4						
		5	2		6		8



As an example, we might have, for our initial list:

- EAT = AT
- ATE = A
- LATER = LOW
- PAN = PILLOW
- CARP = ME.

From these we can derive, for example,

LAP = LEAP

**Anniversary edition**  
Today The Telegraph celebrates 160 years

**Wanted 2015**  
Track 10 pages of our best news coverage

**Feelings?**  
It could be love stories to Marie

## The Daily Telegraph

### Greek crisis ignites as banks shut

**Cameron: now the fightback begins**

**EAT = AT**

**ATE = A**

**LATER = LOW**

**PAN = PILLOW**

**CARP = ME.**

**LAP = LEAP**

# Institute Plan

# Team



James Tagg



Erik Viirre



New Physics



Sir Roger Penrose OM FRS



Stuart Hameroff,



Corporate Affairs

# Vision

To understand the human mind and its place in the cosmos  
inspired by the scientific philosophy of Roger Penrose and to  
use this understanding to the benefit of all.

# Our Collaborative Approach

The World Wide Web was developed as a way to bring researchers together working on large-scale complex problems at CERN. We believe new technologies such as virtual reality and collaborative communications tools can usher in new and more efficient ways of bringing researchers together. It will be an objective of the Institute to experiment with these technologies and innovate in the realm of collaboration. This will involve us building physical and virtual collaboration spaces at our cooperating research institutions.



Slack



GitHub



Basecamp



Oculus



# Institute Structure & Funding

We are seeking funding to scale an institute in a series of 5 phases:-

- A feasibility study grant of \$800,000
- The commitment of a founding grant of \$100m of which
  - The first initial \$10m as a planning/ grant to build core team, infrastructure and put a detailed technical roadmap in place
- Application for government grants and broader funding
- Follow on funding from the primary grant giver and others to secure the institute long term

The Institute will be a non-profit entity in the USA and a registered charity in the United Kingdom along with associated enterprises and future spin outs.

Two associated enterprises are identified at present. One is to develop artificial intelligence-human intelligence collaboration technology and the other a consciousness probe and improved anesthesia.

Separate presentations explains these.

# 5 Years On

In 5 years time what would you hope the institute has achieved?

- Developed a unique collaboration structure that breaks down the normal definition of an institute.
- Developed a good computational understanding of creativity and boundary between human intelligence and artificial intelligence if, of course, such a boundary truly exists. An open question at this time.
- Have a good road map for a computational model of a human neuron with potentially quantum speedup from the newly developing field of quantum computing.
- Made progress in implementing this roadmap
- The ability to probe real neurons in vitro and inform the computational model while allowing real world testing.

What might go 'wrong'?

- Exotic quantum effects may not be significant to consciousness and it might simply be a classical phenomenon. However, quantum modelling WILL be important for neurons and most of the work we do will be very relevant to the practical operation of the human brain, creativity and disease processes.

# Why San Diego?

Physics, Medicine, Math, Nanotechnology, Super Computing, Protein Modelling in one place

California Stem Cell Institute

Human Longevity Inc.

Craig Venter Institute

Scripps

Qualcomm Institute

Salk

Arthur C Clark Center

UCSD

Math, Physics - Non Computability

Neuroscience (#1, ESV)

Cognitive Science (#1, ESV)

Engineering

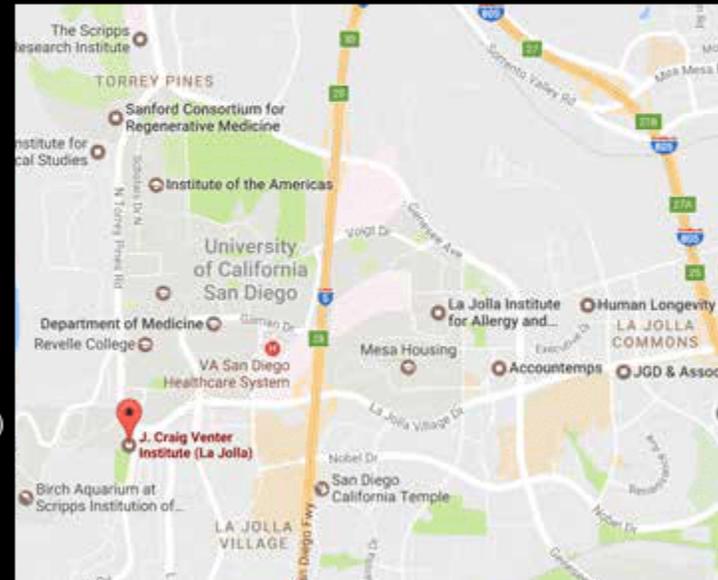
NeuroImaging (#1)

NeuroEngineering (#1, ESV)

Bioengineering (#1 ESV)

Near Google AI and QM, Biotech et. al.

#1, indicates world No1 in world ranking



# Sample Institutes in the Field

Name	Location
<a href="#">Arthur C. Clarke Center for Human Imagination - University of ...</a>	San Diego
<a href="#">Brain and Creativity Institute (BCI) is a research unit of the College of Letters, Arts and Sciences at the University of Southern California</a>	Los Angeles
<a href="#">Institute for Advanced Study</a>	Princeton
<a href="#">J. Craig Venter Institute</a>	San Diego
<a href="#">Kavli Institute for Brain and Mind at University of California, San Diego</a>	San Diego
<a href="#">Perimeter Institute</a>	Canada
<a href="#">Salk Institute for Biological Studies</a>	San Diego
<a href="#">San Diego Biomedical Research Institute: Home</a>	San Diego
<a href="#">The Alan Turing Institute: Home</a>	UK
<a href="#">The Scripps Research Institute</a>	San Diego
<a href="#">Kavli Institute for Fundamental Neuroscience (Kavli IFN)</a>	San Francisco

# FAQ

- *Q: Why is there a need for a new Institute and how are you different?*
  - The Penrose Institute tackles a range of topics from an innovative viewpoint, using newly available quantum technologies. It has a unique perspective on the operation of the human brain and its creative capabilities, and seeks a cross functional integrated approach to understanding this.
- *Q: Isn't the Allen Institute (or others) doing this already?*
  - Other Institutes accept a complex model for neuronal firing and that diseases targets the interior working of neurons, but do not ascribe the level of importance to sub-neuronal structure in the brain, particularly its role in computation. Therefore the work we do is unique. As a corollary it is also controversial with some scientists.

# FAQ

*Q: Why is an entrepreneur founding this Institute rather than a neurologist with a PhD?*

The Institute is being founded jointly by an entrepreneur and a neurologist with a PhD. Putting people with different backgrounds together is a principle of The Institute.

*Q: Hasn't the Gödel proposal put forward by Roger Penrose been discredited?*

No, the proposal is hard to prove. It requires us to know what rules are present in the human brain and show they have been extended. The Institute depends far more on the work of Alan Turing which came later than Gödel, and is a more powerful proof. (Gödel's theorem can be derived from Turing's Halting proof but not vice versa)

*Q: What happens to the Institute if quantum effects in the brain are disproved?*

All chemistry is quantum, so our brains are definitely quantum mechanical structures. The debate is whether 'strange' quantum effects, such as entanglement and non-locality, are in play. Our work on modelling uses quantum principles and will at least allow a better model of the workings of neurons. At best we may prove that some 'strange' quantum behaviors are involved in the brain leading to consciousness. The outcome of the research will contribute to the fields of quantum gravity, cosmology and relativistic quantum information. New insights in the overlap of quantum theory and relativity will enable the development of new technologies.

# FAQ

- *Q: Is The Penrose Institute raising money for a physical building?*
  - Not at this time. We are experimenting with virtual buildings and collaboration spaces, and once this program is well developed we would consider a custom built space or spaces.
- *Q: Do researchers need to work on Orch OR, Twister theory and non-computability to work at The Penrose Institute?*
  - No, The Penrose Institute is inspired by the interests and approach of Sir Roger Penrose across a range of scientific areas concerning the brain, quantum gravity and cosmology. If you have a truly valid, creative and innovative approach to tackling these problems we welcome your involvement. We do believe there is something special to discover about the working of the human brain and we suspect this has implications on our understanding of quantum mechanics.

END

# Appendices

# Creativity and Non- Computability

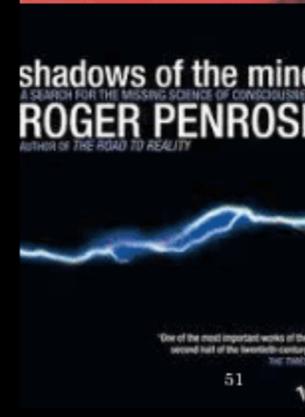
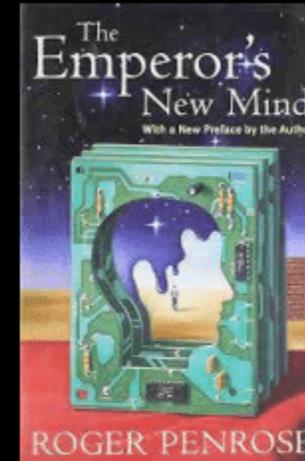
# Is Creativity Computation?

In 1961, J.R. Lucas published "Minds, Machines and Gödel," and in 1989 Penrose published "The Emperor's New Mind", and then later "Shadows of the Mind."

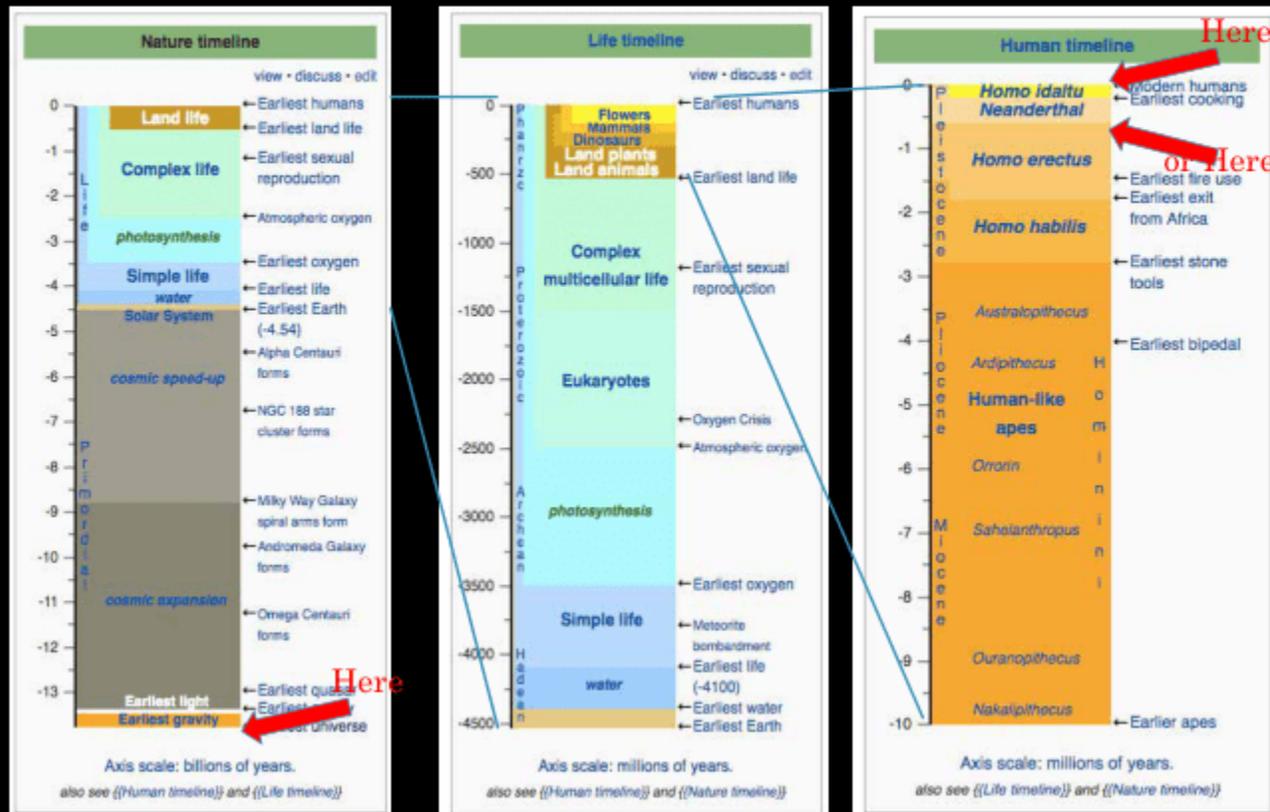
The Lucas-Penrose argument is that Gödel's Incompleteness Theorem states, given any consistent, computable set of axioms, there are true statements about the integers that can never be proved from those axioms. They argue that humans discover these truths and are therefore not axiomatic computable systems.

The problem with this line of reasoning is it requires us to prove what is in a mathematician's brain and demonstrate that the axioms were extended in order to make a new mathematical discovery. Unfortunately there is no way to read out the contents of a given brain.

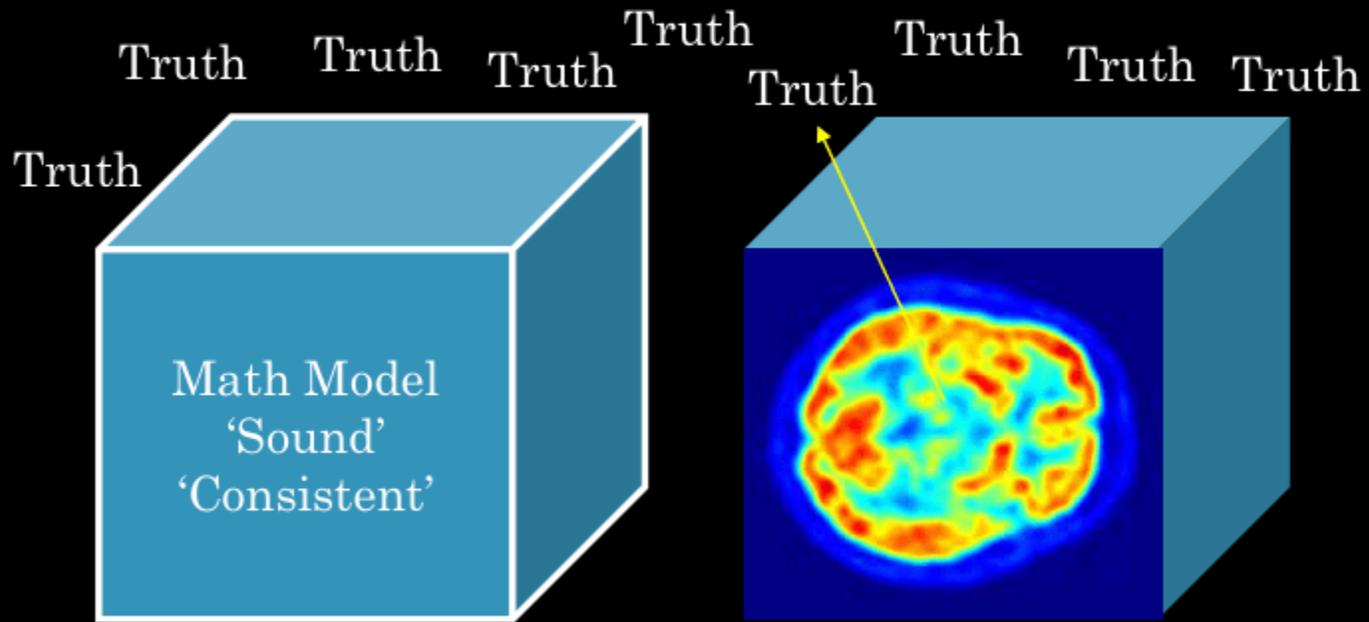
All is not lost as an appeal to Turing's Halting problem proof can be used to show that humans appear to solve non-computable problems.



# When did Creativity Emerge



# Lucas-Penrose Argument



Problem with the Gödel Argument is how do I know what is in the box.



# The Understanding Puzzle

Another way to look at the unique processing method of the human brain is to examine problems that are hard for artificial intelligence. When presented with a problem, the human mind seems to 'understand' the nature of the problem, whereas a computer applies brute force calculation.



*Takahashi/ChessBase*

Chess blocking problems such as this are hard for a computer, but even novice human chess players can understand them. Often a computer will work away at the problem or even blunder and lose the position. Humans 'see' that there is no solution in principle, and that the position is a draw.

Word Puzzles  
Permitted Substitutions

C -> D

A -> O

T -> G

Does CAT = DOG



Humans do it at sight

Computers Calculate

CAT = DAT

CAT = COT

CAT = CAG

DAT = DAG

COT = DOT

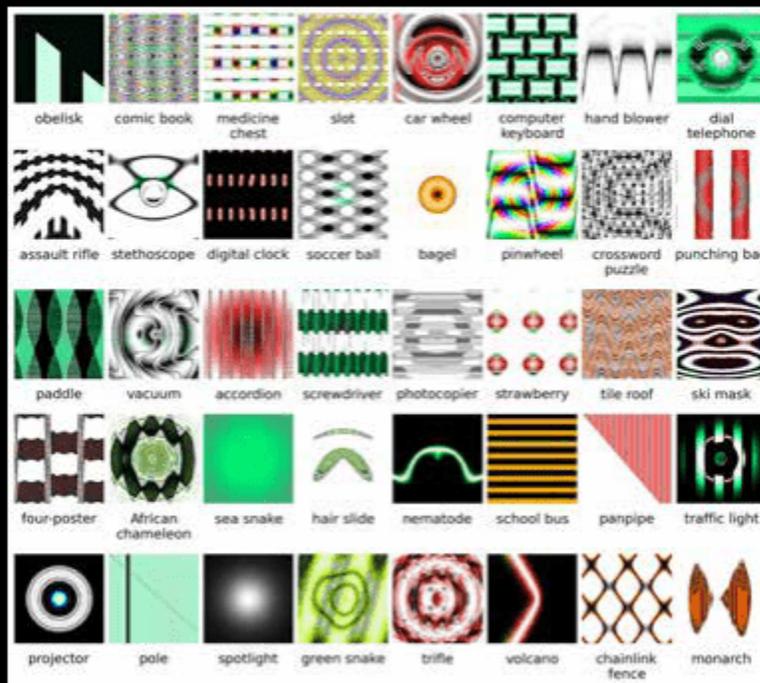
DOT = DOG



YES CAT = DOG

In word substitution puzzles humans appear to solve the problems intuitively whereas a computer will work methodically through all the permutations. It is an open question whether humans are subconsciously computing or if an AI could be programmed to think intuitively.

# Another Example: Image Recognition



All these pictures were misinterpreted by computer image recognition software. Humans don't make this sort of mistake, although it is argued we make other sorts of mistakes and are subject to other failings which do not plague computer systems.

Are brains simply better image recognition computers or do we think differently?

© The Penrose Institute 2016-17

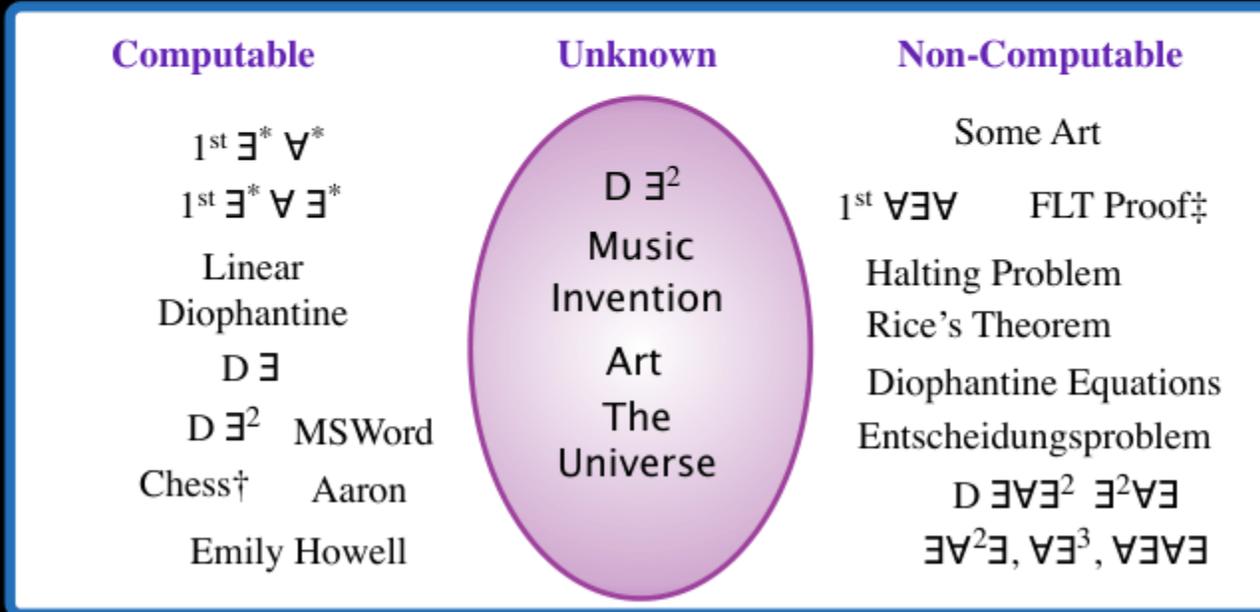
Nguyen A, Yosinski J, Clune J. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. In Computer Vision and Pattern Recognition (CVPR '15), IEEE, 2015. 56

# Human Intelligence and Artificial Intelligence

Artificial Intelligence (AI) has been dramatically improved in recent years, with abilities even superseding human abilities in actions such as visual recognition and complex game play. Further, many AI actions now occur “unsupervised”, with no human action modulating the “training” of the AI system. However, Human intelligence still lies beyond two important boundaries: energy and non-computability. Human and other nervous systems can carry out actions of learning and acting with infinitesimal amounts of energy: a few molecules of ATP. Modern AI requires vast data centers. However, even with infinite digital computing resources (and infinite energy) digital computers, *by definition*, cannot create non-computable results that the human mind can. Mathematical proofs are examples of NC items. An AI could call up the solution to Fermat’s Last Theorem (created by a human) from a memory store, but No AI *can create the solution*. Thus the modern computing complex of AI research will not be able to complete its task: replicating human intelligence. We argue it should not. A new physics will ultimately determine the mechanism of the human brain. In the meantime, understanding of NC mathematics and brain operations will result in a hybrid, that will feature and accentuate the best of both: further freeing Humans from computation and enabling creativity to mix with vast computing power for new and better cognitive actions.

# Non-Computable Classes

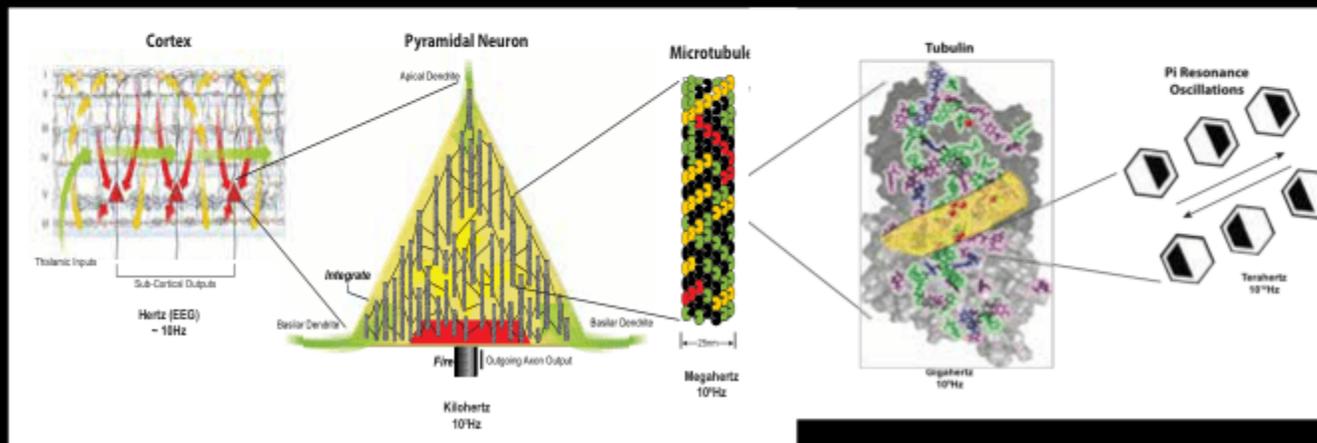
Problems divide into computable and non-computable classes. In the case of a non-computable class there is no general algorithm which will solve all members. This limits the capability of AIs because an AI (being an algorithm) can not find the solution to arbitrary problems within the class, otherwise it would be able to computationally solve any problem in the class, solving the Halting Problem.



# Neuronal Structure

# Neurons Structure & Microtubules

Neurons contain considerable internal substructure which is conventionally thought to provide physical support and energy transport functions. However, some types of neuron have mixed polarity microtubule arrays which have a broken structure unsuited to either of these two functions. With anesthesiologist Stuart Hameroff, Penrose has proposed a theory of consciousness based on non-computable, quantum gravitational computation in these microtubule arrays. The theory is called 'orchestrated objective reduction', 'Orch OR'. Despite being widely attacked when first proposed the theory has several merits over other quantum theories, in particular it is experimentally verifiable. It is both possible to test whether collapse occurs at certain energy levels and to see whether non-computable computation is going on - i.e. whether neuron arrays can 'beat' a Turing Machine when performing certain tasks.



# Molecular Action of Neurons

Neurons are structures built of regular proteins, membranes, water and dissolved ions.

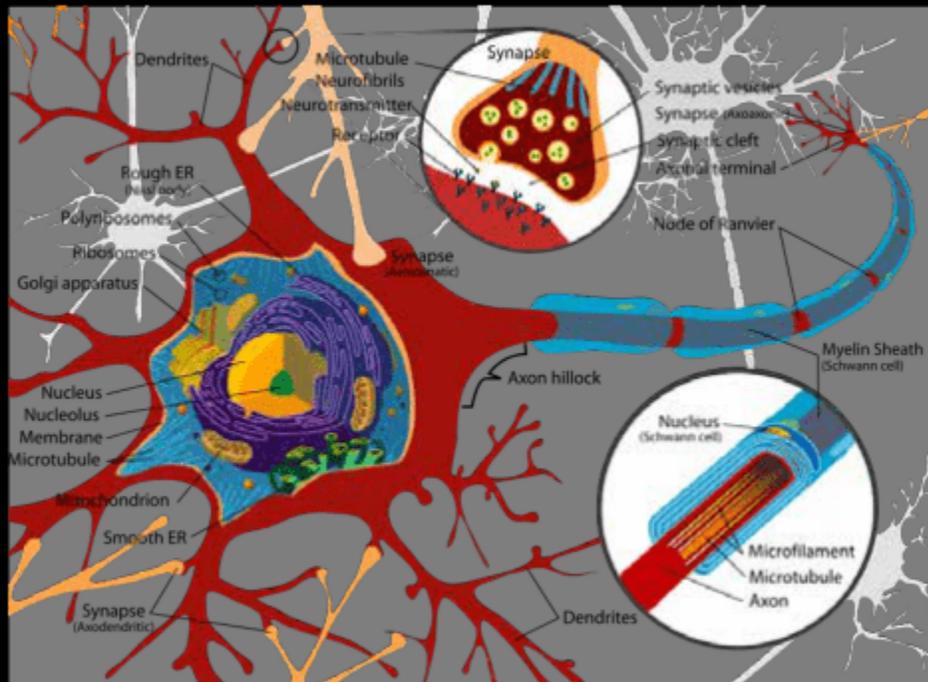
Their activity is physical-chemical action on those structures: the biology of life and thought

A crucial part of the mission of the Penrose Institute is to understand the molecular structure of neurons and the physics of actions occurring on them.

Combining observational physics and physical chemical modelling will be done:

- Modelling water and then dissolved ions inside and outside neurons
- Modelling ion channels and e-m physics beyond Hodgkin-Huxley
- Measurement and modelling of microtubules
- A physical model of the complete axon with activities expected by prediction and verified by experiment
- A quantum mechanical examination of micro-tubular activities
- Application of the quantum biological principles of life, such as those coming from botany, to look for energy and computational efficiency
- A quantum biological model of neurons

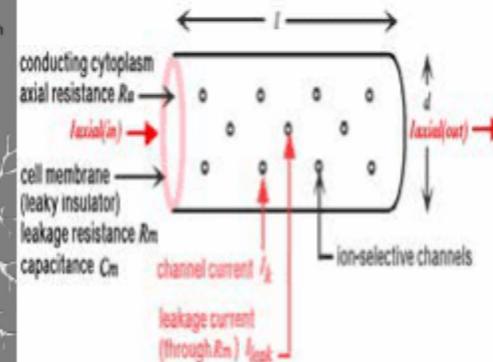
# Neuronal Structure



## Hodgkin-Huxley

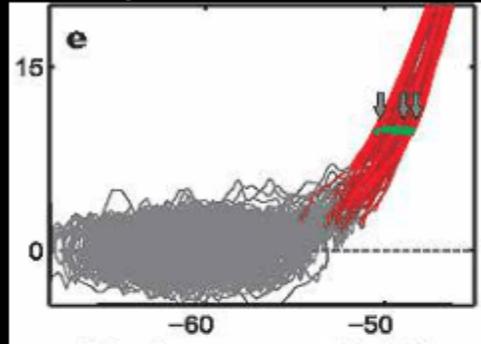
### Neuronal Firing Model

The standard model of neuronal firing is as a classical avalanche effect kicked off by the initial synaptic firing of adjacent neurons, but experiments show the firing does not accurately follow this model.



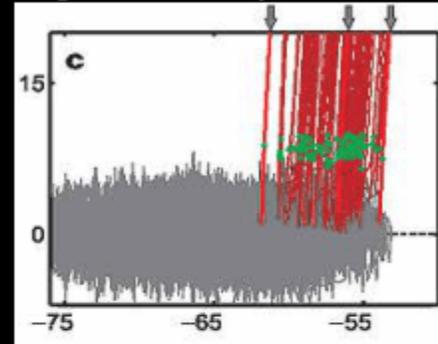
# Experimental Deviation from Predicted Hodgkin-Huxley Behavior

Projected from H-H

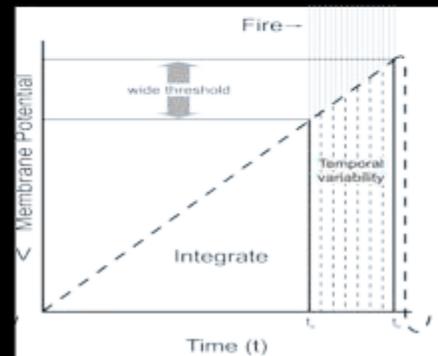
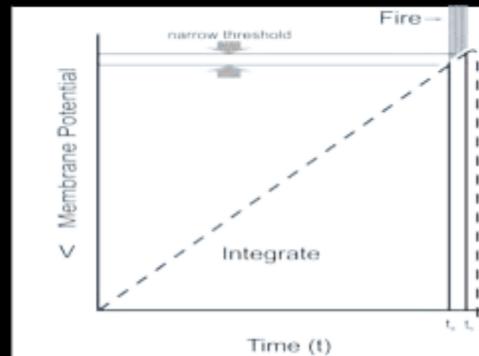


Membrane potential

Experimentally Observed



Membrane potential

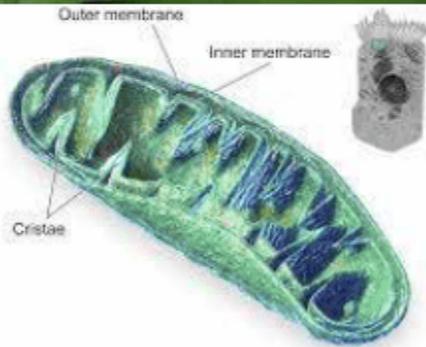


Naundorf B, Wolf F, Volgushev M. Unique features of action potential initiation in cortical neurons. *Nature*. 2006 440(7087):1060-3.

© The Penrose Institute 2016-17

63

# Models of Human Microtubules show Potential Quantum Behavior



**Mitochondria**

