THE IMPOSSIBILITY OF ARTIFICIAL IMAGINATION

A Question of Genuine Consciousness

HUY TRINH

IS IT POSSIBLE TO CREATE GENUINE CONSCIOUSNESS, ARTIFICIAL INTELLIGENCE? THIS PAPER WILL EXAMINE THIS QUESTION THROUGH THE PHILOSOPHICAL UNDERSTANDING OF IMAGINATION. IMAGINATION IS THE CREATIVE FACULTY THAT NOT ONLY SEPARATES MAN FROM BEAST BUT ALSO ALLOWS MAN TO RECOGNIZE HIMSELF, AND IT HELPS HIM TO SEE BEYOND HIS IMMEDIATE BEING. WITHOUT IMAGINATION, MAN IS BLIND TO FUTURE POSSIBILITY AND UNABLE TO CONCEIVE OF FUTURE EVENTS. IMAGINATION PRODUCES MAN'S MENTAL COGNITION THAT MAKES HIM A RATIONAL AND CREATIVE BEING. THE TWO MODERN METHODS OF A.I. PROGRAMMING ARE INCAPABLE OFIMBUING THEIR ALGORITHM WITH IMAGINATION. WITHOUT IMAGINATION, A.I. NOT ONLY LACKS GENUINE CREATIVITY, BUT IT ALSO RISKS NEVER UNDERSTANDING SELF-AUTONOMY AND THEREFORE COULD NEVER BE A FREE AND INDEPENDENT ENTITY.

INTRODUCTION

For millennia, the power of creation has been the prerogative of omnipotent divinities. With the advent of computer and artificial intelligence (A.I.), man is slowly replacing the gods as the givers of consciousness and intelligence. Before science fiction writers can fantasize about A.I.'s plan for global domination; however, computer scientists must answer one of the most important questions about A.I.: "Can A.I. imagine?" And the answer is no. The ability to imagine denotes a consciousness that sees beyond itself and recognizes the infinite possibilities of the world. Without this faculty, any self-professed intelligence cannot qualify as a truly independent conscious being. The two current approaches to A.I programming—Formalism and Bottom-Up—are incapable of duplicating the creative or reflective power of imagination. Neither is it conceivable that programmers will come up with an algorithm that can represent imagination. The logician Kurt Gödel's incompleteness theorem demonstrates the inherent weakness of computer algorithm and its confining boundary. The algorithm of A.I. creates an objective mental frame that prevents A.I. from being subjective in its judgment or from going beyond its immediate environment. This structural limitation prevents A.I. from being greater than its program, from being self-aware, and from plotting world domination. Imagination is a unique creative element of human consciousness that acts both as a mirror for self-reflection and a telescope to see beyond the narrowed vision of the present. It is impossible to mathematically transfer this subjective productive element to an objective structure.

THE ROLE OF IMAGINATION

To understand why A.I. cannot imagine, the questions of what imagination does and why A.I. needs imagination in order to be conscious must be addressed. Man's ability to comprehend the sublimity of images, transcending colors and dimensions, depends on imagination. The German philosopher Immanuel Kant credits imagination for "bringing together the manifold of intuition" of the mind so mere images and thoughts can become a source of cognition.[i] By combining sensual experiences with reason and emotions to create images and ideas that might not be there empirically, imagination unifies sensation and abstract thoughts together giving the mind a medium to comprehend the physical world.[ii]

These reproductive and productive aspects of imagination have earned it the ire of ancient philosophers like Plato and Augustine. They attack imagination not only for its mimicking but also for its corruption of the real. Plato denounces imagination's irrationality and its users—i.e. artists, poets—as corrupters of reality.[iii] They are mistaken, however, on the crucial role the irrationality of imagination plays in cognition.

"Imagination is a productive and creative faculty because it is not constricted by facts and empirical actuality."

Imagination "releases thing[s] from their contingent status as facts and grants them an ideal status as possibilities . . . "[iv] Imagination is a productive and creative faculty because it is not constricted by facts and empirical actuality. Imagination's irrational nature frees it from rationality's predictability. If reason was the sole cognitive faculty then reactions to external stimuli would be solely logical and predictable without variations and changes. Because of its irrational nature, imagination can see beyond the real and into what is possible or impossible beyond the fact: the *"irreality,"* and the most important of such *irrealities* is the future—a mental world of tomorrow. Imagination channels many different mental faculties—e.g. memory, reason, emotion and impulses—when projecting the images of the possible. And imagination is con-

tinuously active in formulating and projecting futuristic reality throughout a person's life. Every individual wakes up each morning and prepares for the day's activities in order to satisfy future needs; the individual eats to keep the future self alive, sleeps to keep the person awake the next morning. The past is gone and the present is fleeting, but only through the future, with its "irreality" and infinite possibilities, can humans find the reason for continuous survival—even if those dreams are hollow and meaningless.[v] This reproduction of reality by the imagination can be detected in brain-scans: when the brain imagines, it mimics the actual world in its imagination. One neurobiologist documents that "mental activity in the absence of non-mental stimuli can result in similar if not identical neural activity as that generated by the material stimuli."[vi] One can see that the human brain not only reproduces the empirical world through imagination, but it also behaves the same way in this imaginative world as it would have in the real one.

The futuristic possibilities conjured by imagination are the primary components in the recognition of the self and its freewill. Imagination, by not being a subject to external imagery, can see the essence, the platonic form of the object without being bound to sensorial boundary. Imagination can "engage in a free play [or freedom to create] since no definite concept restricts [it] to a particular rule of cognition."[vii] Because of this freedom to process, imagination can examine the mental selfhood of a being outside of the being.[viii] Kant explains that the "I" which thinks does not exist under its own prerogative; instead, it is rooted in the synthesis between the infinite possibilities and the actual reality via the productive imagination.[ix] The "I" in Descartes' famous philosophical maxim, "I think therefore I am," denotes a being that recognizes its mental selfhood. To recognize its mental selfhood, the mind must be able to separate itself from the rest of reality and put itself in a place independent of its body; imagination is this haven that gives the mind the freedom to reflect about itself. It is only through the realization of its mental selfhood that a conscious person recognizes that he has freewill. From the abstraction of its mental selfhood, the "I" uses reason to

TO BE SELF-AWARE, A.I. MUST ACKNOWLEDGE ITS "I" AS AN AUTONOMOUS, FREE-THINKING BEING.

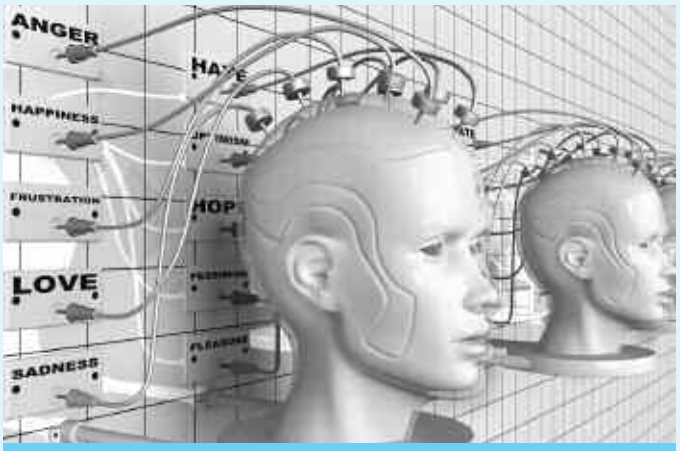
identify its complete autonomy within its consciousness. Imagination's Delphian vision expands the autonomy of the mental self through its formulation of futuristic possibilities. The concept of free will emerges from this mental autonomy, and with imagination, the mind decides which futuristic possibilities it wants to enact. We can conclude that man's mental existence, his "I," is an existence via imagination and reason—they are like lines joined together to create a complete plane of cognition; without either one to connect to the other, the plane is neither complete nor whole.

Intelligence experts agree that for an A.I. to be called a conscious being it must be greater than the sum of its parts—streams of meaningless digital characters. To be self-aware, an A.I. must acknowledge its "I" as an autonomous being, free from the orders its creator had placed on it; it must also match the human consciousness at both the sublime and irrational levels. With irrational imagination, A.I. can envision the infinite possibilities that are beyond its experience and program. If the A.I. is bound by the rationality of its program, then it is trapped in a cage it does not know exists. Imagination and mental freedom are inseparable; if devoid of imagination, then A.I. will be nothing more than a mirror mimicking the world as is and not as the world might be.

LIMITATIONS OF FORMALISM

There are two schools of artificial intelligence: *Formalist A.I.* and *Computational Intelligence (Bottom-Up)*. The former tries to program a mind into the machine and the latter tries to develop a mind through education; the difference between these two approaches is similar to that of empiricists and rationalists. Nevertheless, neither school can reproduce the imaginative consciousness due to a shortcoming they both share: their dependency on algorithm.

Formalism believes mental activity is "the carrying out of some well-defined sequence of operations [algorithms],"[x] and consciousness is a "passive concomitant of possession of a sufficiently elaborate control system and does not, in itself, actually 'do' anything."[xi] In other words, the human mind is a complicated set of mathematical sequences, which can be discovered and duplicated. Consciousness is the side effect of the neural network working seamlessly together to form the illusion of "oneness." The Formalist's theory of cognition is similar to the platonic form or Descartes' innate ideas. By programming complex logical rules and premises, Formalist A.I. researchers believe a system of consciousness will emerge from their mathematical algorithm. One of the early researchers of A.I., the mathematician and logician Alan Turing, outlines the basic principle of cognition in his hypothetical Turing Machine—a machine that runs on an unimaginably large and comprehensive algorithm that takes into account every conceivable potential input and output. The Turing Machine's only function is analyzing the data from the input and then releasing an output by following the preprogrammed instruction. Both the input and the output are codified into numerical forms making it easier to compute; the outputs are then translated from numbers into human language.[xii] Most A.I. supporters see the human mind as nature's finest and most complicated Turing Machine; the mind takes in data through sensations and then reacts according to the circumstances. Formalist researchers see chess-playing computers as a prime example of what is possible with the Turing Machine; their machine can out "think" the best human chess player when programmed with the algorithm of chess.[xiii]



ARTIFICIAL INTELLIGENCE IS CHAINED TO SETS OF ALGORITHMS, ELIMINATING ANY ELEMENT OF GUESS WORK OR IRRATIONALITY

Despite its promises, Formalist A.I. suffers from many debilitating drawbacks. One of the drawbacks of Formalist A.I. is its consumption of huge amounts of computer power and the vast amount of computer codes involved, which also increases the chance of mistake and bugs. Formalist researchers acknowledge the computer's *serial calculation*—one calculation at a time—is vastly inferior to the *parallel distributed processing* (PDP) power of the human brain—multiple calculations performed simultaneously by the same computational unit. Nevertheless, Formalist researchers are confident that a newer and much faster supercomputer will overcome the speed limitation. Formalist researchers' greatest hope is the quantum computer—a super computer that has overcome the limitation of quantum mechanics to compute at an unimaginable speed. Quantum computers, researchers believe, can replicate the "oneness" of consciousness that serial computers can never reproduce. The physicist John Hopfield ignited

a new revolution in computer thinking with his groundbreaking discovery of spinning atoms behaving like neurons when organized into a neural-like network. This discovery promises a new phase in computer's evolution, one that might match the computing power of the brain within a few decades. The only remaining, but nevertheless most daunting, obstacle is the complicated neural networking and algorithm that simulate intelligence. Formalist researchers contend that this challenge will be tremendously difficult but not impossible.[xiv]

Beside computational speed, the Formalist approach faces numerous problems that are not only technologically daunting but are also logically impossible to overcome. First, the algorithm that runs a Turing Machine is, at its most basic level, a series of material conditional—if-then statements. Since a serial computer is coded in algorithms, it cannot have one flawed instruction; addtionally, each axiom is built from the previous one so a mistake would create errors throughout the algorithm. The brain contains many unreliable elements that can break down when stressed; evolution, however, has "[exploited] the fault tolerance and functional persistence that PDP automatically confers ."[xv] Since each neuron performs only a small margin of the overall work and the job it performs can be handled by another neuron, even if a tenth of all the neurons were to disappear, the overall performance would only suffer marginal degradation. In contrast, "[Formalist] machines are limited," the philosopher Paul Churchland points out, "that is, to computing mathematical functions whose inputs and outputs can be expressed as ratios of *whole numbers*."[xvi] This allegory means that the inherently rigid and inflexible structure of algorithm must have all of its mathematical sequences working harmoniously to produce a rational result, represented by whole numbers. If any input or output contradicts its mathematical sequence or produces an answer that it does not recognize (e.g. an imaginary number that has no real existence), the result will be erroneous. Basically, Formalist A.I. cannot be irrational by going against its program. As a whole, Formalist A.I. is vastly inferior to the human mind in term of flexibility and reliability.

The next obstacle Formalist researchers must face is the feasibility of the programming process. Creating a massive algorithm requires a large number of programmers working collectively for a long period of time and the constant updating of the algorithm to confront new challenges. "The universal Turing Machine is a mimic" of its programmers; it is neither autonomous nor independent from the programmers' instruction.[xvii] The Formalist A.I. can be sophisticated and complicated, but it can never evolve on its own, supplementing and expanding on its older algorithm to the extent that it can be free from its own algorithmic limitations. In contrast with the Formalist machine, human's mental judgment depends upon "complicated interconnected-combination of sense-data, reasoning and *guesswork*."[xviii] This guesswork represents the non-algorithmic imagination that gives humans the freedom to not only do what is right and logically sound but also to be irrational and fatuous. The structure of Formalist A.I. is the antithesis of imagination because it leaves no room for self-improvement or self-discovery. The A.I. simply cannot defy its rule of cognition, exercise free play or guess. Ultimately, A.I. is prisoner to its algorithmic cage. The ultimate challenge to a Formalist researcher is the problem of algorithm, and it is this limitation that prohibits the creation of genuine A.I.. Before examining this dilemma further, an investigation into the second approach to A.I. is necessary.

THE BOTTOM-UP APPROACH

The second school of A.I. programming, known as computational intelligence (C.I.), or Bottom-Up, approaches A.I. from an empirical direction. Bottom-Up researchers abandon the cumbersome and complex algorithm for a naturalistic intelligence. Their method resembles Locke's view of cognition where the mind is a "tabula rasa," ready to be filled with new experiences and knowledge. Unlike their Formalist counterparts, Bottom-Up researchers use robots in their experiments. Through trials and errors, these ro-

bots learn from their mistakes and build on their progress, slowly edging toward intelligence. A Bottom-Up robot is not preprogrammed with restricted instructions; rather, it has the freedom to experience the world as it is. The robot is equipped with sensual instruments and a pain-pleasure index to record its experiences in a realistic fashion. Although most of the Bottom-Up robots learn at an incredibly slow speed (partially due to their small sizes and low computational power), they appear to genuinely learn from what they are experiencing. From learning how to hold forks to recognizing human faces, these robots gather, store, and respond to data. Researchers hope that with the increase in computer power, they can build a smaller, more efficient robot.[xix] Unfortunately, despite its promising start, the Bottom-Up school faces structural limitations of its own and shares the same algorithmic set-back as that of Formalist A.I.

A Bottom-Up robot gathers data differently than a Formalist A.I.; however, it only replaces the programmer with natural experiences. The robot still operates under an algorithm—one that expands as it learns—therefore, it has all the shortcomings of an algorithmic machine. The robot's most severe setback is its dependency on empirical stimulus: it only cognizes what it has experienced; everything beyond its memory is incomprehensible. The robot has no abstract concept of the objects it perceives. It breaks down and stores the data it has received without interpreting that data as anything more than data. Moreover, as a reactive machine, the robot cannot grasp the temporal dimension. It cannot take a leap of faith and perform an action that it has not experienced before; therefore, it is forever trapped in the present. The phenomenological theory of imagination recognizes this shortcoming of perception without imagination:

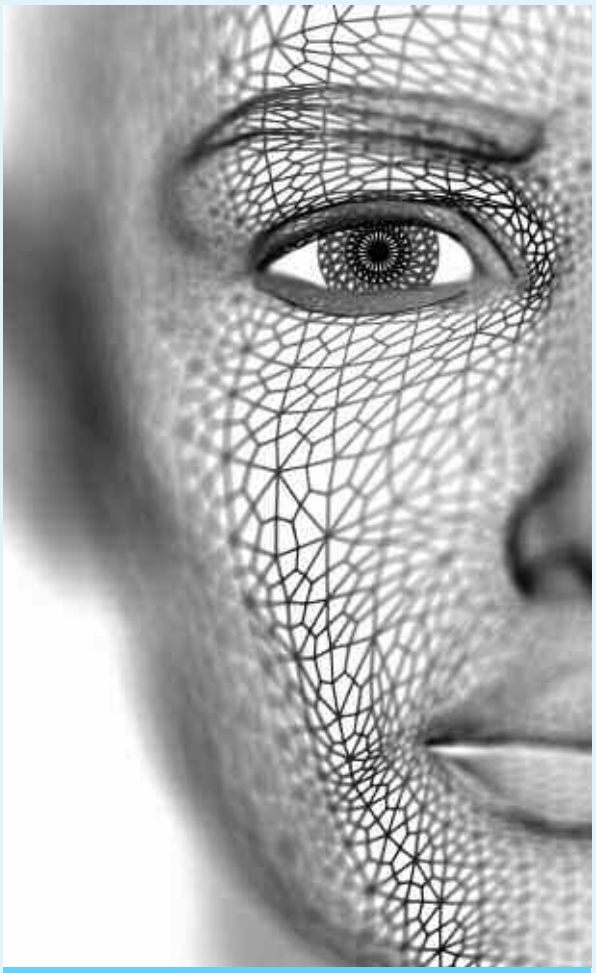
> *From perception alone we could never conclude that extension is integral to thinghood as an eidetic necessity. For if we were confined to our immediate perception we could not imagine instances where this might not be so, and thereby (by process of negation) discover it must be so.*[xx]

The absence of *a priori* imagination prevents the robot from ever developing an ego to comprehend the "thinghood" of the objects it has perceived. To remediate this problem, researchers of Bottom-Up must create artificial imagination. Their only alternative is to program an *a priori* imagination into their robot using a set of algorithm. This is where researchers of Bottom-Up confront the same impossible challenge as researchers of Formalist A.I.; both schools of artificial intelligence have met an obstacle they cannot bypass: programming a conscious algorithm.

Both the Bottom-Up and Formalist A.I. researchers' faith in innovations in computational technology is a gamble that will not pay off. Technological advancement in hardware does not mean researchers can produce the software to run it. Computer hardware might be able to simulate the firing process of neurons and adapt the brain's PDP system, thus improving the structural makeup of the machine, increasing its speed and efficiency, but the computer's basic software remains the same. A serial or a parallel computer will always remain a Turing Machine that needs a sophisticated algorithm to run it. A.I.'s last hope, the quantum computer, is also an unlikely source of intelligence. Despite its unfathomable speed, quantum computers cannot perform "non-algorithmic operation."[xxi] The physicist Hopfield was only suggesting that "intelligence *might* arise from the quantum theory of mindless atom, without any programs whatsoever," hinting at how our own intelligence might have evolved but does not promise that it can duplicate the intelligence through programmatic means.[xxii] In the end, a quantum computer is only a very fast Turing Machine. The final case against A.I. not only destroys the hope of building an autonomous A.I. through algorithm, but it will also provide evidence supporting the non-algorithmic nature of imagination and human consciousness.

GÖDEL'S PROOF AGAINST A.I.'S IMAGINATION

The final case against genuine A.I. consciousness comes from the Austrian logician Kurt Gödel and his incompleteness theorem, which states that "no finite set of algorithmic



WHILE A.I. CAN BECOME INCREASINGLY SIMILAR TO THE HUMAN RACE, IT CANNOT ACHIEVE TRUE HUMANITY

procedures can generate all arithmetic truths." In other words, no system of algorithm or mathematical equation can prove itself infallible under its own rule without risking a paradox. Because each algorithm needs to have another algorithm to validate itself, no algorithm is ever complete. Since Formalist A.I. is built from algorithms, it can never prove any propositions or ideas that it invents or discovers on its own using its own algorithm.[xxiii] This severe limitation means that A.I. will be far short of full consciousness. The human mind, however, faces no such restriction thanks to its imaginative faculty. The infinite possibilities conjured by imagination allow the mind to sidestep any mental restriction placed on it; this means that the mind is capable of answering the questions it raises, e.g. a mathematician can discover an unproved theorem and then prove his own theorem. Formalist researchers counter by arguing that the mind runs a very complex algorithm that has both questions and answers within it. Ironically, the argument raised by A.I. supporters creates a paradox when examined more closely. If A.I. researchers claim to have discovered the algorithm to the human mind, they must present a proof that it is indeed the right algorithm. According to Gödel's indisputable theorem, however, an algorithm cannot examine its own system without fallacy; therefore, the human mind cannot safely examine the algorithm of human consciousness, and if researchers cannot validate this algorithm, than they cannot assert that it is the correct algorithm.[xxiv] This paradox dispels both the possibility of ever knowing the algorithm of human consciousness and the possibility of its duplication. The mathematician Roger Penrose concludes that by convincing "ourselves of the validity of Gödel's theorem, we not only 'see' it, but, by so doing, we reveal the very non-algorithmic nature of the 'seeing' process itself."[xxv] Through Gödel, one can see the mind operates in a non-algorithmic fashion. The mind does not have an algorithm because imagination allows consciousness to "emancipate itself from its immersion in the world of actuality and to return to a realm of pure possibility."[xxvi] An algorithm can never imagine an improbable idea or understand that reality is an "infinite metamorphosis"; it only sees the world that it was designed to see.

CONCLUSION

Imagination is "a power capable of intending the unreal as if it were real, the absent as if it were the present, the possible as if it were actual."[xxvii] Imagination is a surreal and mystical faculty of our consciousness. Descartes once questioned the nature of his reality because he did not trust his imagination; he fears that reality is a fantastic dream created by his imagination. Descartes is right in his fear because there is nothing imagination cannot touch, no

boundary that it cannot cross. With imagination, reality becomes a blur; without imagination, reality has no meaning: it is because of this indefinable faculty that the human mind remains unique. The flaw with A.I. designers is that they "cared relatively little about exactly what processes take place inside us, so long as they implement the right input-output function."[xxviii] Artificial Intelligence is a product of a rational system; therefore, it could never act contrary to its purpose. It cannot be programmed to create the unreal out of the real—doing so would create a counter system to its present system. Lacking an independent mental selfhood, A.I. is incapable of distinguishing the right from the wrong system. Intangible abstractions like time and "thinghood" require irrationality as well as rationality. Unfortunately, these two properties are counterintuitive to the A.I.'s system. An intelligent machine operates under a mathematical language that cannot comprehend the meaning behind the symbols that represent its being. It is human who assigns the meta-mathematical language to the inherently meaningless binary.[xxix] In the end, it is human consciousness that interprets the essence behind A.I. Without our imaginative mind to give it purpose and meaning, A.I. will only exist in a state of reactive behaviors.

It is doubtless that one day in the near future, humans will create a life-like artificial intelligence that will mimic our exact behaviors. It will speak with a life-like voice, behave like our dearest of companions, and fight like the most patriotic soldier. It will perform its tasks with a level of efficiency we could never match. We will relegate more and more tasks to our A.I., becoming more dependent on its services than ever before. Some will claim that we have become pets to our own creation. Despite the important roles A.I. will play in our society, the human mind will remain the sole genuine consciousness on this planet. The secret of our consciousness remains locked within the Pandora's box of our imagination. It is this imagination that has liberated us from an empirical prison; however, when we look beyond ourselves, we will realize that we are alone in our freedom.

ENDNOTES

i. Kearney and Rasmussen (12)
ii. Kearney 1998 (48)
iii. Kearney 2001 (93)
iv. Kearney 1998 (19)
v. Kearney 1998 (23)
vi. Rose (219)
vii. Kearny 2001 (12)
viii. Kearny 2001 (21-22)
ix. Kearney 2001 (49)
x. Penrose (17)
xi. Penrose (405)
xii. Penrose (7)
xiii. Kaku (74-5)
xiv. Kaku (83-84)
xv. Churchland (13)
xvi. Churchland (243)
xvii. Penrose (51)
xviii. Penrose (99)
xix. Kaku (87-88)
xx. Kearney 1998 (25)
xxi. Penrose (399-402)
xxii. Kaku (83)
xxiii. Nagel and Newman (111)
xxiv. Penrose (417)
xxv. Penrose (418)
xxvi. Kearney 1998 (22)
xxvii. Kearney Poetic (16)
xxviii. Churchland (251)
xxix. Churchland (244)

REFERENCES

Churchland, Paul M. 2000. *The Engine of Reason, the Seat of the Soul.* Cambridge: MIT Press.

Kaku, Michio. 1997. *Visions: How Science Will Revolutionize the 21st Century.* New York: Anchor Books.

Kearney, Richard. 1998. *Poetics of Imagining.* New York: Fordham University Press.

Kearney, Richard. 2001. *The Wake of Imagination.* London and New York: Routledge.

Ed. Kearney, Richard and David Rasmussen. 2001. *Continental Aesthetic: Romanticism to Postmodernism.* Malden: Blackwell Publisher.

Penrose, Roger. 1989. *The Emperor's New Mind.* New York: Oxford University Press.

Nagel, Ernest and James Newman. 2001. *Gödel's Proof.* Ed. Douglas Hofstadter. New York: New York University Press.

Rose, Steven. 2005. *The Future of the Brain: the Promise and Perils of Tomorrow's Neuroscience.* New York: Oxford University Press.