Mental Experience and the Turing Test: This Double Face is the Face of Mathematics

Dedicated to Alan Turing's 100th Birthday

Leon N Cooper*

Departments of Physics and Neurosciences
Institute for Brain and Neural Systems
Brown University
Providence, RI, 02912

and

Sorin Istrail**

Department of Computer Science
Center for Computational Molecular Biology
Brown University
Providence, RI, 02912

We seem to be on the threshold of understanding the physiological basis for learning and for memory storage. But how would knowledge of events on the molecular and cellular level relate to human thought? Is complex mental behavior a large system property of the enormous numbers of units that are the brain? How is it that consciousness arises as a property of a very complex physical system? Undoubtedly, these questions are fundamental for a theory of the mind. On the other hand, there are questions of basic importance, pioneered by Turing, for his theory of the "human computer," that is, discrete state machines that "imitate" perfectly the mental processes achievable by his "human computer"; we will refer to it, although the name is only partially true to his vision, as his theory of "computational intelligence." Finding even a level of commonality to discuss both a theory of the mind with a theory of computational intelligence has been one of the grand challenges for mathematical, computational and physical sciences. The large volume of literature, following Turing's seminal work, about the computer and the brain and involving some of the greatest scientists of all time is a testimony to his genius.

In this paper we discuss, in the context of the Turing test, recent developments in physics, computer science, and molecular biology at the confluence of the above two theories, inspired by two seminal questions asked by Turing. First, about the physical not reducible to computation: "Are there components of the brain mechanism not reducible to computation?" or more specifically, "Is the physical space-time of quantum mechanical process, with its so called Heisenberg uncertainty principle, compatible with a [Turing] machine model?" Second, about computing time: "[in the Turing test] To my mind this time factor is the one question which will involve all the real technical difficulty." We relate the above questions to our work, respectively, on superconductivity and quantum mechanics, and the Ising model and the proof of its computational intractability (NP-completeness) in every 3D model, and share lessons learned to discourage, under high and long-term frustration of failure, the retreat under the cover of the positivist philosophy or other evasions.

Inspired by von Neumann, we relate Turing's questions and his test difficulties to von Neumann's thesis about the "peculiar duplicity" of mathematics with respect to the empirical sciences. As von Neumann put

it: “This double face is the face of mathematics.” With a non-*a priori* concept of truth “the very concept of ‘absolute’ mathematical rigor is not immutable. The variability of the concept of rigor shows that something else besides mathematical abstraction must enter into the makeup of mathematics... something nonmathematical, somehow connected with the empirical sciences or with philosophy or both, does enter essentially…”
----

1. The Turing Test
2. Turing’s Questions
3. Turing’s Physical vs. Computational Question and Theory of Mind
4. Turing’s Time Factor Question and Computational Intractability
5. Von Neumann’s Double Face of Mathematics
6. Charging on

1. The Turing Test

In the 1950 paper “Computing Machinery and Intelligence,” Turing proposed what was later coined “the Turing test.” In the paper he asked:

“*Can machines think? I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. I further believe that no useful purpose is served by concealing these beliefs*.” [Computing Machinery and Intelligence]

Clearly Turing’s prophecy that “machines will eventually compete with men in all purely intellectual fields” came true by the end of his century. In 2010, IBM’s Watson won its *Jeopardy* match against the television show’s all-time human champions, and Apple iPhone’s Siri “intelligence” makes you feel that “Promethean irreverence” that Turing referred to in his “Intelligent Machinery” paper. If Prometheus stole fire from Zeus and gave it to mortals, we now see “intelligence” stolen from mortals and given to machines!

Two of the common confusions in reading Turing’s paper were pointed out by Chomsky, giving his fair renditions of Turing’s views. “It would be a confusion to seek ‘empirical evidence’ for or against the conclusion that brains or machines understand English or play chess; say, by resort to some performance criterion. This seems a fair rendition of Turing’s view. … The great artisan Jacques de Vaucanson did not seek to fool his audience into believing that his mechanical duck was digesting food, but rather to learn something about living things by construction of models, as is standard in the sciences. Turing’s intentions seem similar in this regard.”[1]

What follows is a discussion on both of these topics, the role of the empirical in mathematical modeling as well as physics and computer science aspects of the models.

2. Turing’s Questions

Turing’s seminal work provided the much-needed formal definition of a computational procedure, and therefore, a mechanical proof in a formal logical system. Godel in 1964 praised his achievements:

---

[1] Noam Chomsky, “Turing on the “Imitation Game,” in *The Turing Test*, S. Shieber ed., pp. 317-321, MIT Press, 2004

"Due to A. M. Turing's work, a precise and unquestionable adequate definition of the general concept of formal system can now be given. … Turing's work gives an analysis of the concept of 'mechanical procedure' (alias 'algorithm' or 'computation procedure' or 'finite combinatorial procedure'). … A formal system can simply be defined to be any mechanical procedure for producing formulas, called provable formulas."

Turing asked a number of fundamental questions, which also shows his philosophical thinking about the mathematical and physical modeling for the brain. His work was clearly aligned with a dichotomy: "the computational" (or "mechanical thought") is defined by the Turing machine to be rooted in finiteness of the mechanical procedure's number of steps, and "the physical" -- although, as his biographer A. Hodges notes[2], only one sentence in all his work asks about the physical not reducible to computational. To have further insight into this separation, concepts such as consciousness and free will were not, in his view – expressed especially after his post-war work -- phenomena belonging to the computational component.

T1. "*What is the physical mechanism of the brain*?" The brain mechanism he thought of as a finite set of rules, doing one step at a time.

T2. "*Is the physical space-time of quantum mechanical processes, with its so-called Heisenberg uncertainty principle, compatible with a [Turing] machine model?*" The action of the brain must be computable – he adds "quantum mechanics" in his 1951 BBC radio talk presenting his famous 1950 paper "Computing machines and intelligence" in Mind, 49, 433-460. His biographer, A. Hodges, comments that "this is the *only* sentence in all of Turing's work that points to something physical that may not be reducible to computable action."

Turing's zest for the hardest problems shows again, in his wish list, to find approximations for quantum mechanics, which although modeled then and now mathematically within a space with infinitely many dimensions – he wanted a new quantum mechanics theory with finitely many dimensions (about 100) aiming at making quantum mechanics fully predictable and perhaps more finite.

Turing machines are rooted in the concept of "doing one step at a time," but here also the quantum mechanics' wave-function vis a vis macroscopic observation enters the picture.

T3. "*Are there components of the brain mechanism not reducible to computation*?"

T4. *"Brain behavior should in principle be predictable by calculatio*n" although computational simulation, as is usually done with a concept of pseudo-randomness, was not adequate; he asked,*"What does computational 'simulation' mean?"* He hinted at the physics definition of "random" instead*.* He was *"wondering whether any machine can do as much as the brain."*

"This machine [the brain] should be the sort whose behavior is in principle predictable by calculation," Turing said. " *We do not know how any such calculation should be done and Sir Arthur Eddington on the account of indeterminacy in quantum mechanics claims that no such prediction is even theoretically possible.* "

T5. The time factor in the Turing test. "*To my mind this time factor is the one question which will involve all the real technical difficulty.*"

3. Turing's Physical vs. Computational Question and Theory of Mind

---

[2] A. Hodges, *The Enigma*

It has been obvious to many of us for quite a while that the brain is only marginally a computing system. It is no more designed for logic or reason than the hand is designed to play the piano. If it is designed at all, the design concerns survival and, in an ordered world, survival is enhanced by rapid (even if occasionally incorrect) decision-making.

Through years of painful education, somehow, our brain has achieved the ability to reason and execute rules as well as to associate. (It seems clear that association is much easier and more natural than reasoning – the preserve of some of our more cerebral types.[3]) We say somehow because it is not yet all worked out (there are, of course, major areas of current research directed toward determining precisely how visual and other information processing are achieved) but, perhaps conceptually, it is there.

Having designed such a system (its practical and commercial value aside), do we have a machine (a generalized Turing machine?) that thinks? In effect, this is the question the Turing test is designed to evade – an evasion due, in part, to an excess of positivism – a fear of and/or aversion to mentalism or assumptions about the internal workings of the mind that cannot be directly verified by experience. It is an evasion that, in our opinion, is totally contrary to the nature and purpose of scientific thinking.

Successful science has given us just such machines (actual or conceptual) that work behind the actual events. The greatest include Newton's laws, Maxwell's equations and Schrodinger's equation. Such entities as molecules and/or atoms were assumed to exist (an assumption that was vigorously contested in the 19th century with positivistic-type arguments) long before they were "seen." It is not necessarily the case that every element of the "behind the scenes" machinery can be directly observed. In quantum mechanics, for example, the wave function is not directly observable. The consequences of this sometimes invisible machinery can, however, be put into correspondence with experience. The essence of the positivist argument (as actually employed by Einstein and Heisenberg) is not that we cannot introduce entities that are not directly observable but, rather, that if an entity is not observable (e.g. absolute time in special relativity or simultaneous position and momentum in the quantum theory) it need not appear in theory.

Thus a satisfactory theory of mind not only is allowed but, in our opinion, requires the introduction of mental entities. We will be satisfied only when we see before us constructs that can have mental experience, when we see how they work, how they come about from more primitive entities such as neurons.

All of this can be summarized by saying that the Turing test is not sufficient. At best, it provides a perfectly responding "black box" but no knowledge of how it all works – and that, in our opinion, is just the knowledge we want.[4]

What is more surprising, perhaps, is that for what is hardest to understand, the origin of the complex of mental experience: consciousness, awareness of ourselves, feeling, passing the Turing test is not necessary. Nothing in the perfect machine response gives us any indication of how much experience comes about. Or

---

[3] Among the glories of human intellectual achievement is, from the hazy associations that are the natural capacity of the brain, just this creation and successful application to messy real-world situations of precise language (to make what is said depend on what was said before) and logical reasoning. In a *fin de siècle* copout, a California update of dialectical materialism, Hegel come to Los Angeles, we witness a seasonal twist of intellectual fashion that heralds fuzzy logic (a sometimes useful engineering tool), recently become fuzzy thinking as more appropriate than Aristotelean logic to describe the less than sharp boundaries of the real word. Thus it is proposed that we convert the razor-sharp distinctions of the trained and athletic mind to the indistinct muttering and imprecise groping of the intellectual couch potato, that to the complexities of the world we add the

[4] We could, of course, take the black box apart to investigate the sequence of instructions that gave us all of the right answers. It remains an open question whether this would necessarily shed any light on the phenomenon of mental activity.

in the words of the late philosopher Hans Jonas, "The capacity for feeling, which arose in all organisms, is the mother value of all."

The deepest error and what may be most misleading is the attempt to equate the activity of the brain with a reasoning system. A more appropriate view is that of Albert Szent-Gyorgi:

"The brain is not an organ of thinking but an organ of survival, like claws and fangs. It is made in such a way as to make us accept as truth that which is only advantage. It is an exceptional, almost pathological constitution one has, if one follows thoughts logically through, regardless of consequences. Such people make martyrs, apostles, or scientist, and mostly end on the stake, or in a chair, electric or academic."

For the understanding of mental behavior as consciousness or feeling, it is thus clear that the Turing test is not only not necessary but is totally irrelevant. One can reason, or at least perform logical operations, without feeling: Mechanical calculators do it all the time. Dogs and cats (even turtles, probably) feel but can't answer many questions.[5]

Let us then put this famous and ingenious criterion to rest and confront again the underlying problem. Can we understand the human mind (all of its components: reasoning, feeling, self-awareness) and its presumed origin in that biological organ, the brain? Could we, in the extreme, construct a machine that was conscious? (Whether a machine in this sense could be a Turing machine we will leave for others to answer.) What are the steps required so that a machine (algorithmic or not) can experience mental activity? The non sequitur "How would we know?" is an evasion. (How do we know anything?) Whether we can be sure that another creature and/or a machine is conscious is independent of the understanding of how it is that consciousness arises as a property of a very complex physical system. This, not reasoning power, is the profoundest mystery surrounding that biological entity, the brain.

What has made this problem even more perplexing is the confusion of the various components of mental activity, the failure to distinguish what we believe we know – at least in principle – from what we do not yet understand. Although it is true that we are still far from understanding, for example, how the visual system processes and sorts information, such questions can be precisely formulated and it seems reasonable to believe that answers can be constructed from materials available. The same might be said for reasoning (logical and otherwise). But what is the source of our mental experience – our conscious awareness?

Mental awareness or consciousness themselves have many components, some easy to understand, some still incomprehensible: an on-off switch; memory in storage, memory in play; the distinction between ourselves and the external world (still the subject of vast philosophical cerebration) are, we believe, easily understood.

In Turing's work the computational and the physical visit each other and this dance is at the heart of the synergies of the computational with the theory of mind. Von Neumann's concept of "complication" although in need of development of a full theory of it gives a computational theorem about the emergence of a systems property at a certain level of complexity of organization, with obvious similarities with mental entities, which he is able to define formally in the context of another biological process of high complication: self-reproduction. The ending paragraph from his "The General and Logical Theory of Automata" charges the course of a systematic theory of automata and the threshold between degenerative and self-supporting.

---

[5] In Turing's defense (if he needs a defense), we should note that such mental behavior was not what he had in mind. His concern was the output of human reasoning – not human feeling.

"All these are very crude steps are in the direction of a systematic theory of automata. They represent, in addition, only one particular direction. That is, as I indicted before, the direction towards forming a rigorous concept of what constitutes 'complication.' They illustrate that 'complication' on its lower levels is probably degenerative, that is, that every automaton that can produce other automata will only be able to produce less complicated ones. There is, however, a certain minimum level where this degenerative characteristic ceases to be universal. At this point, automata, which can reproduce themselves, or even construct higher entities, become possible. This fact, that complication, as well as organization, below a certain minimum level is degenerative, and beyond that level can become self-supporting and even increasing, will play an important role in any future theory of the subject." [The General and Logical Theory of Automata, 1948]

3.2 The regulatory genome and the computer: molecular mechanisms for "memory"

The Regulatory Genome. Knowledge of the molecular biology mechanisms of cell regulation, especially genomic cis-regulatory systems and gene regulatory networks, governing all cells, is now available, revealing the cell regulatory mechanism – and for those cells of the nervous system they are key parts of the "neural mechanism." Our paper, "The regulatory genome and the computer," with Eric Davidson of California Institute of Technology, [19] the foremost experiment biologist in gene regulatory networks and the regulatory genome, is written in the same compare-and-contrast format as "The Computer and the Brain," as a homage to von Neumann's last book written in large measure on his deathbed. In it, we present a first comprehensive view of the information processing capability of the genomic regulatory system of the cell. Davidson's experimental work is focused exclusively on *causality,* as the exquisite genomic regulatory mechanisms, locked down by evolution, can only be revealed through experimental DNA perturbations. In this respect, for the biological cell, Davidson's work stands out as the flagship work on causality-based systems, as articulated by Einstein in 1953. "Development of Western science is based on two great achievements: the invention of the formal logical system (in Euclidean geometry) by the Greek philosophers, and the discovery of the possibility to find out causal relationships by systematic experiment (during Renaissance)." Our paper brought together a full information processing system view, building on our decade-long collaboration focused on genomics, logic functions of the genomic cis-regulatory code, and transcriptomics.

In short, there are many thousand of cis-regulatory modules -- DNA regions upstream of genes a few thousand bases long -- in animal genomes that are "wired" together into very large networks that control biological processes of such development. Each cis-regulatory module is an "information processor" and each gene is controlled by several cis-modules; then the genes and their cis-modules are assembled in gene regulatory networks. Each cis-module has "inputs" which are transcription factor proteins that bind to short DNA subsequences of the cis-module. The communication of information is done by means of diffusion of transcription factors as opposed to pre-organized wires in the electronic computer. The design principles are dramatically different in respect to time, speed, synchrony-asynchrony, memory, hardware and software, parallel computing processors, as well as fascinating new concepts of fault-tolerance.

Von Neumann: "Not the language of mathematics but the language of the brain." Without question mathematics is the foundation of the design of our electronic computers -- from PCs and cell phones to supercomputers, a.k.a. von-Neumann architecture computers. The information processing system of the cell shows that the logical elegance, of mathematical language nature, for the world of the Turing-von Neumann electronic computers is totally different than the world of the "genomic computer." If the electronic computer is relying on a decisive separation of key design components and concepts such as hardware-software, time-memory, synchrony-asynchrony, sequential-parallel, rigid-evolving, discrete-continuous, what characterizes the cell information processing is exactly the lack of such separation, in fact, an intertwining of the dualities, with dramatically different concepts where separation boundaries are no longer well-defined. A most relevant bottleneck for all electronic computers, the so-called "von Neumann bottleneck," subject to intense research to overcome it in future designs, with no solution in hand yet, is elegantly resolved in the biological world so information processing there is free of it.

4. Turing's Time Factor Question and Computational Intractability

Turing observed, reflecting of the "Turing test" technical difficulties, that computational complexity, or how he put it, "the time factor", is where "**all** real technical difficulty is" (emphasis is ours). *To my mind this time factor is the one question which will involve all the real technical difficulty.*"

Computational complexity is now a major research area of computer science, achieving millennium stability (to paraphrase von Neumann's comment on Geometry which achieved "bimillennial stability") with the key concept, NP-completeness, and the problem P=?NP recognized as one of the seven Millenium Problems – unresolved problems of the highest diificulty that will shape the mathematical landscape of the 21st century.

Turing would have liked the concept of NP-completeness – computational problems that apparently have only exponential algorithms for their solutions, which means for regular size of instances, the solutions is requiring very large time, of scales of millions of years, not available to human timelines. He would have liked to estimate and approximate those answers. Like his "Turing second large number" that he estimated: the "Shakespeare" number N was such that 1/N is the probability that "this piece of chalk will jump from my hand and write a line of Shakespeare on the board before falling to the ground."

One of the most celebrated NP-complete problems, the problem of obtaining analytical closed form solutions for the three-dimensional Ising Model, has this property for every 3D model.[6] The relationship of the time factor in the Turing test and computational intractability is a theme of much reflection and requires a separate analysis.

5. Von Neumann's "This Double Face is the Face of Mathematics"

Von Neumann presented his thesis on the duality and inseparability of mathematics as empirical and non-empirical science. "The most vitally characteristic fact about mathematics is, in my opinion, its quite peculiar relationship to the natural sciences. ... In modern empirical sciences it has become more and more a major criterion of success whether they have become accessible to the mathematical method or to the near-mathematical methods of physics. Indeed, throughout the natural sciences an unbroken chain of successive pseudomorphoses, all of them pressing toward mathematics, and almost identified with the idea of scientific progress, has become more and more evident. Biology becomes increasingly pervaded by chemistry and physics, chemistry by experimental and theoretical physics, and physics by very mathematical forms of theoretical physics. ... This double face is the face of mathematics, and I do not believe that any simplified, unitarian view of the thing is possible without sacrificing the essence."[7]

---

[6] S. Istrail, Statistical Mechanics, Three-Dimensionality and NP-Completeness: I. Universality of Intractability of the Partition Functions of the Ising Model Across Non-Planar Lattices, Proceedings of the 32nd ACM Symposium on the Theory of Computing (STOC00), ACM Press, p. 87-96, Portland, Oregon, May 21-23, 2000

[7] "The Mathematician," John von Neumann 1947

> The difficulties with the Turing test could be seen at the heart of such duplicity, and attempts at disentangling the computational and the physical of the brain models would have a tremendous challenge to face as much of it would come from this duplicitous inseparable unitarian view.[8]

As von Neumann warned us:

"... the very concept of 'absolute' mathematical rigor is not immutable. The variability of the concept of rigor shows that something else besides mathematical abstraction must enter into the makeup of mathematics. ... Something nonmathematical, somehow connected with the empirical sciences or with philosophy or both, does enter essentially…"

6. Charging On

These questions are sufficiently difficult so that we have been subjected to the usual evasions – Cartesian dualisms: variations of homunculus proposals; solutions of one mystery by invoking another: consciousness arises in the quantum measurement process or where gravity meets quantum theory; total refusal to confront the issue: consciousness arises "somehow" when a machine executes the proper algorithmic processes; total retreat under the cover of positivist philosophy: how would we know if a machine were conscious… and so on.

We have heard such arguments before. They seem to be typical responses to the frustration of failure in attacking really difficult scientific problems. First try and fail. Follow this by proving a solution is impossible or irrelevant. Toy with the notion that a new law of nature is involved. Then, when the solution is found, complain that it is really trivial or (even better) that it was suggested in some obscure comment once made in a paper you published a long time ago.

On a personal note , one of us (L. N. C) experienced all of this in the course of developing a theory of superconductivity, [9] also a complex and subtle consequence of an interacting many-component system, in this case the quantum mechanics of electrons in a metal. After the fact, one rather well-known physicist expressed his disappointment that "such a striking phenomenon as superconductivity [was]… nothing more exciting than a footling small interaction between electrons and lattice vibrations" – thus missing the point in operatic style.

It could turn out that we must invoke a new "law of nature": pour the conscious substance into the machine – a position not unfriendly to a common view of mankind in its emerging years. But the conservative scientific position is to attempt to construct this seemingly new and surely very subtle property from the materials available – those given to us by physicists, chemists and biologists (as has been done many times before: celestial from earthly material, organic from inorganic substances, the concept of temperature from the motions of molecules – and so on). If this cannot be done (perhaps one could be patient enough to give us a couple of years to try) then we will genuinely have made one of the profoundest discoveries in the history of thought – consequences of which would shape and alter our conception of ourselves in the deepest way.[10]

---

[8] "The Mathematician," John von Neumann 1947

[9] Microscopic Quantum Interference Effects in the Theory of Superconductivity- Nobel Lecture, December 11, 1972; http://www.nobelprize.org/nobel_prizes/physics/laureates/1972/cooper-lecture.html

[10] A distinction must be made between a new assumption, an entirely new entity (the equivalent of the addition of Euclid's fifth or parallel axiom to the first four that distinguishes between Euclidean and non-Euclidean geometries) and an unexpected and non-inevitable construction from the materials available (e.g. living creatures from chemicals or novels from letters). These latter are often highly dependent on initial conditions. Even temperature, a seemingly straightforward construction from kinetic theory and statistical mechanics, requires equilibrium systems.

In another personal experience, concerning the provability of NP-completeness of the Ising model for every 3D model, one of us (S. I.) observed and recorded the despairs expressed by the greats of physics and mathematics when explicit analytical formulas were not available, despite the invention of new methods of extraordinary depth by a dream team of scientists working on it – Feynman, Fermi, Fisher, Temperley, Kac, Onsager (our former Brown University colleague, and Nobelist as well).[11] Well, computational complexity as defined by NP-completeness, is at a more fundamental of the fundamental number of steps required for a computational solution in the Turing sense, for which no computational method can evade. [12]

Here is a glance at the helplessness feelings of that era:

"The three-dimensional case does exhibit a phase transition but exact calculation of its properties has proved hopelessly difficult. The two-dimensional case ... was solved by Lars Onsger in 1944. Onsager's solution, a veritable 'tour de force' of mathematical ingenuity and inventiveness, uncovered a number of surprising features and started a series of investigations, which continue to this day. The solution was difficult ... and George Uhlenbeck urged me to simplify it. 'Make it human' was the way he put it. ... Even Feynman got into the act. He attended two lectures I gave in 1952 at Caltech and came with the clearest and sharpest formulation of what was needed to fill the gap. The only time I have ever seen Feynman take notes was during the two lectures. Usually, he is miles ahead of the speaker but following combinatorial arguments is difficult for all mortals." -- Mark Kac 1968

The prospect that such a program could be carried out elicits occasionally paranoid reactions: cries of reductionism or, as expressed by John Lucas of Oxford, "the rubbishing of human experience."

As for reductionism, we have always been somewhat mystified as to what the fuss is about. Scientists (as mentioned above) have been constructing seemingly new and elevated entities from base material as part of their daily exertions since Thales showed us the way. When such a construction cannot be made, something new must be added. There seem to be no shortage of voices advising us to add that something new before we have had a reasonable chance to construct from the old. We advise patience, along with the reminder that our brain is programmed to jump to conclusions.

The rubbishing of human experience is a greater concern since no small number of the *nouvelle vague* computer and robotic types seem only too happy to do just that. To us this appears to be a reflection of the joy they have experienced and the wisdom they have gained in their passage through this vale of tears. The value we place on our own experience is something we determine ourselves and would never (we hope) forfeit to any machine or in fact (except to a limited extent) to anyone else. This value is completely independent and totally unaffected by any "reductionist" explanation of how our mental activity comes about. No more than a detailed knowledge of the chemistry of digestion affects our appreciation of the product of a great chef or the bouquet of a fine wine.

Well then how, from materials made available to us by unrepentant reductionists: electrons, protons, atoms, molecules, DNA, RNA, receptors, enzymes, proteins, membranes, neurons, axons, dendrites, synapses …, can we construct an entity that has mental experience?

---

[11] S. Istrail, Statistical Mechanics, Three-Dimensionality and NP-Completeness: I. Universality of Intractability of the Partition Functions of the Ising Model Across Non-Planar Lattices, Proceedings of the 32nd ACM Symposium on the Theory of Computing (STOC00), ACM Press, p. 87-96, Portland, Oregon, May 21-23, 2000

[12] S. Istrail, *Statistical Mechanics, Three-Dimensionality and NP-Completeness: I. Universality of Intractability of the Partition Functions of the Ising Model Across Non-Planar Lattices*, **Proceedings of the 32nd ACM Symposium on the Theory of Computing (STOC00)**, ACM Press, p. 87-96, Portland, Oregon, May 21-23, 2000

We believe that there is a reasonable evolutionary sequence leading from sunshine, lightning and a reducing atmosphere to molecules, the primitive protein soup and to more and more complex structures. Even the simplest cells show reflex-like chemically directed responses to various stimuli (aversion or attraction). As has been said, "protoplasm is irritable." (But how does it come to feel irritable?) It seems reasonable to believe that there is advantage for organisms that can communicate from one end to the other, that electrical communication is a very efficient way of doing this, and that excitable membranes provide the means.

There is further advantage in the innovation of a nervous system that is plastic (the transmitter driven synapse provides an excellent option), for the animal can now learn and store memories of past experience. This animal can adapt to environment changes in less than evolutionary time.

And there is surely advantage (most of the time) in exercising the option of mental experience – feelings of pleasure and pain, awareness of individuality, instilling a directive to "jump to conclusions" to "accept as truth that which is only advantage" as a means of producing life preserving behavior in complex and somewhat unpredictable real-word situations.

All of this can at least be sketched. We can guess (at least conceptually) how from primitive feelings such as pleasure and pain more complex mental states could be constructed. But how the essential primitive feeling arose out of materials such as reflex reactions to hot and cold -- how, somewhere in the distinction between those events that produce physical reactions in ourselves and those that do not, in the interplay of present sensory input with memory of past experience, our self-awareness, our mental experience, our consciousness arose -- remains a deep mystery.

To paraphrase Shannon:

"Can machines feel?" "Sure they can. We're machines. We feel"

But are we? And if so, how?

About the authors

*Leon Cooper is the Thomas J. Watson, Sr. Professor of Science, Professor of Physics and Neurosciences, and Director of the Institute for Brain and Neural Systems at Brown University. In 1972, he received the Nobel Prize in Physics (with J. Bardeen and J.R. Schrieffer) for his studies on the theory of superconductivity completed while still in his 20s. In 1968, he was awarded the Comstock Prize (with J.R. Schrieffer) of the National Academy of Sciences, Descartes Medal of the Academie de Paris, Université Rene Descartes. He holds seven honorary doctorates.

**Sorin Istrail is the Julie Nguyen Brown Professor of Computational and Mathematical Sciences and former Director of the Center for Computational Molecular Biology at Brown University. Before joining Brown, he was the Senior Director of Computational Biology at Celera Genomics where his group played a central role in the construction of the Sequence of the Human Genome. In 2000, he obtained the negative solution (computational intractability) of the three-dimensional Ising Model for every 3D model. He is Professor *Honoris Causa* of the "Alexandru Ioan Cuza" University, Iasi, Romania.